LAND ECONOMICS

a quarterly journal of

PLANNING, HOUSING & PUBLIC UTILITIES

"WHATEVER MAY BE THE LIMITATIONS WHICH TRAMMEL INQUIRY ELSEWHERE
WE BELIEVE THE GREAT STATE UNIVERSITY OF WISCONSIN SHOULD EVER
ENCOURAGE THAT CONTINUAL AND FEARLESS SIFTING AND WINNOWING BY
WHICH ALONE THE TRUTH CAN BE FOUND."

From Report of Board of Regents, University of Wisconsin, 1894

SOCIAL RESEARCH

An international quarterly, founded in 1934, published by the

GRADUATE FACULTY OF POLITICAL AND SOCIAL SCIENCE

of the New School for Social Research, New York

Contents for Summer 1959 · *Anniversary Issue*

(Volume 26, Number 2)

Book Reviews

Published in Spring, Summer, Autumn and Winter

Subscription $7.50 a year (Foreign $8.00) — Single copies $2.00 plus postage

66 WEST 12TH STREET, NEW YORK 11, N. Y.

JOURNAL OF FARM ECONOMICS

Published by THE AMERICAN FARM ECONOMIC ASSOCIATION

Editor: Robert L. Clodius

University of Wisconsin, Madison, Wisconsin

VOLUME XLI	AUGUST 1959	Number 3

This journal contains additional articles, notes, book reviews, and announcement of new bulletins in agricultural economics and is published in February, May, August, November and December. Yearly subscription is $9.

Secretary-Treasurer: C. DEL MAR KEARL

Department of Agricultural Economics

Cornell University, Ithaca, New York

CO

Th

Ph

At

Lo

Co

Co

Or

P

A

N

W

I

M
M
m
p
A
iz
in

LAND ECONOMICS

a quarterly journal of

PLANNING, HOUSING & PUBLIC UTILITIES

CONTENTS   AUGUST 1959

The Nonwhite Population Surge to our Cities

By PAUL F. COE*

On May 3, 1948 the United States Supreme Court ruled that residential property covenants dealing with racial restrictions could no longer be enforced in federal or state courts and on June 15, 1953 ruled that alleged damages arising from the violation of racial property covenants were not collectible with court assistance. In furtherance of the principle of equal rights for all, the Court ruled on May 17, 1954 that racial segregation in public schools is now unconstitutional. To implement this ruling several states are presently formulating plans to effectuate public school integration. In view of the kaleidoscopic social changes growing out of these decisions which deal with racial segregation, this article has significant timeliness.

THE tremendous population explosion which is hurling nonwhites out of the rural South and into our big cities is one of the outstanding sociological phenomena of our time.[1] In this rapid nationwide movement, nonwhites are surging into central cities in particular and into standard metropolitan areas (SMA's) in general.[2] This dramatic shift, which has become increasingly pronounced during the last decade and a half, arouses deep concern among city administrators, educators, businessmen, economists, and housing analysts alike. In close accord, these groups perceive dual implications in this trend. First, the nonwhite inmigration introduces new problems of inter-racial adjustment in many northern and western cities where previously these problems had been largely absent. Hence, it generates new imperatives de-

* Economist and Statistician, Federal Housing Administration, member of the Federal Committee on Standard Metropolitan Areas, and of the Technical Committee on Standard Metropolitan Area Definitions, Washington, D. C. The analysis, views, and opinions expressed in this article are the writer's and not necessarily those of any organization or agency.

[1] The term "nonwhite" includes Negroes, Indians, Japanese, Chinese, Philipinos, Koreans, Asiatic Indians, Polynesians, and other Asiatics. Persons of Mexican birth or ancestry who are not definitely Indian or of other nonwhite race are classified as white. Of the 1950 nonwhite continental United States total, 95.5 percent were Negroes, as computed from the *1950 Census of Population* (Washington, D. C.: United States Bureau of the Census), Vol. II, Part 1, Table 35. The Census Bureau defines a respondent as nonwhite if he is considered such in his community, however small his nonwhite mixture may be.

[2] A standard metropolitan area includes the county, or counties, containing a central city or cities of 50,000 or more inhabitants, and any contiguous counties, if, according to specified criteria, they are essentially metropolitan in character and are socially and economically integrated with the central city. In New England slightly different criteria of area integration are required to reflect the particular political structure of that region. The generic definition of an SMA is given more fully in the *1950 Census of Population* (Washington, D. C.: United States Bureau of the Census), Vol. I, p. XXXIII; for a precise delineation of each SMA, see Tables 26 and 27 of that volume.

manding speedy recognition and early solution. Secondly, it highlights the fact that as a growing number and proportion of nonwhites are drawn to the central cities these places tend to become disproportionately the repositories of lower-income population. This results principally from the tendency of lower-income nonwhites to flow centripetally into older and usually cheaper dwellings within the central cities while the higher-income white population contrariwise out-pours centrifugally into newer and generally costlier housing in the suburbs.

The ultimate effects of these conspicuously divergent population vectors on the very fabric of urban society pose a challenge of major dimensions. With so much of the urban complex involved in this process of racial redistribution a city pattern is emerging which differs significantly in social scope from anything which the nation's cities have experienced heretofore. What is the anatomy of this growth pattern? What socio-economic factors are bringing it about? What further consequences can it produce? How can the entangled problems induced by this influx be alleviated? These are a few of the questions which are considered in the following paragraphs.

At the outset it is important to recognize that the special problems involving nonwhites are basically of socio-economic origin. They are sociological because the bulk of the nonwhites still remain submerged beneath the main stream of American life even though numerically many nonwhites are distinguished doctors, lawyers, scientists, educators and community leaders. They are economic because nonwhites as a group are economically disadvantaged as compared with the white component of our society, even though growing numbers of nonwhites are becoming affluent.

Movement to SMA's. The nonwhite population of the United States is gravitating rapidly to the nation's large SMA's. Between 1940 and 1950 the number of nonwhites living in the 168 SMA's rose from 5.72 to 8.25 million.[3] Moreover, nonwhites poured into SMA's at more than twice as fast a rate as did the whites, 44 percent compared with 20 percent. During the same period the number of Negroes and other minority racial groups living outside of these populous centers actually dropped from 7.74 to 7.51 million. Thus, by 1950, 52 percent of the nation's nonwhite population was living in the 168 metropolitan areas. In 1940 the comparable percentage was only 42; and in 1930 it was 39. Furthermore, in 1950 some 7.51 million nonwhites still resided outside SMA's to feed the stream of city-bound nonwhites.

Concentrations in Cities. The most important fact regarding the in-pouring of nonwhites to SMA's is that the bulk of the increase is occurring in the central cities. This is manifest in that, of the 2.53 million nonwhites who moved to or were born in SMA's during the last decennial Census period, 82 percent settled in the central cities while only 18 percent located in the suburbs. Indeed, in these core cities nonwhites increased about five times as fast as did the whites over the past decade, 48 percent compared with 10 percent. As the SMA's grow, it is the white population which tends to spill over into the suburbs; 69 percent of the SMA white increase was in the suburbs and only 31 percent in the central cities.

While in most SMA's the increasing number and proportion of nonwhites

[3] For a treatment of nonwhite population trends in the United States from 1790 to 1950 and for an exhaustive statistical analysis of nonwhite shifts to SMA's and cities from 1940 to 1950, see Paul F. Coe, "Nonwhite Population Increases in Metropolitan Areas," *Journal of the American Statistical Association,* June 1955, pp. 283-308.

in central cities are accompanied by a numerical increase in white persons, the central cities of 22 SMA's experienced an actual decline in white population over the past decade. For example, Chicago gained 227,000 nonwhites and lost a net of 3,000 whites. In Cleveland, nonwhites increased by about 65,000 while whites decreased by 28,000. St. Louis gained 45,000 nonwhites and lost over 4,000 whites. In Newark the non-whites increased by 29,000; whites decreased by 20,000. Pittsburgh, Buffalo, Jersey City, Nashville, Youngstown, and Trenton also are numbered among the 22 that gained nonwhites and lost white population.[4] Moreover, in addition to the net white decreases noted, a number equal to all of the white natural population increase attributable to the excess of births over deaths in those 22 cities likewise out-migrated to the suburbs. This constituted a substantial additional out-flow of white population, for during this period natural increase reached a high level. Furthermore, much of the natural increase in the white population of many other central cities also out-migrated to the suburbs even though there remained a net increase of white persons in those central cities.

Current Trends. Evidence that the accumulation of nonwhites in central cities continues right up to the present is afforded by an examination of the special enumerations conducted by the Bureau of the Census from 1955 through 1958 in selected places of 50,000 or more population and with 1,000 or more non-whites. In two-thirds of the 42 cities meeting these size criteria and having

special censuses, nonwhites increased much faster than the white rate and than the nonwhite national rate of 22 percent from 1950 to July 1958.[5] Parenthetically, the white national rate of increase was only 14 percent during the same period. The median increase in the 42 cities was 37.5 percent for nonwhites compared with only 9.4 percent for the all-race total. Indeed, nonwhite population grew by over 80 percent in the California cities of Riverside, Santa Ana, San Diego, and San Mateo, in Syracuse, New York, and in Ft. Wayne and Hammond, Indiana. In San Jose nonwhites increased by 115 percent, in San Bernardino by 140 percent, and in Compton, California by a resounding 375 percent (all percentages adjusted downward to discount for any annexation increases.) In each of these cities the white relative increase was far less than the nonwhite. In Syracuse nonwhites almost doubled while whites actually decreased by over 10,000 during the seven-year period. Four of the seven cities where the white rate of increase exceeded the nonwhite are located in the South.

Local estimates indicate that in the nation's capital nonwhites had increased from 35 percent of the all-race total in April 1950 to 48 percent in January 1958.[6] This means that, since 1950, nonwhites living in the city of Washington have increased by over 130,000 as against a drop of about 68,000 in number of whites and as against a nonwhite increase of 96,000 for the entire preceding decade. Should this pace persist, over half of Washington's population will be nonwhite by 1960. The ultimate color

[4] Table, "Nonwhite Population Trends in Standard Metropolitan Areas, in Central Cities, and in Their Suburbs, 1930, 1940, and 1950," (Washington 25, D. C.: Division of Research and Statistics, Federal Housing Administration). This table gives detailed nonwhite population trends for each of the 168 SMA's in the United States in 1950 and was drawn upon in the analytical development of some of the early paragraphs of this paper.

[5] *Current Population Reports* (Washington, D. C.: United States Bureau of the Census), Series P-25, No. 193, Table 2; Series P-28, Nos. 805-1,241; and *1950 Census of Population*, Vol. II, Parts 2-50, Table 33, and Table 36 for southern states.

[6] J. P. Pickard, "Eight Years of Growth and Change in Greater Washington," *Washington Board of Trade News*, October 1958, p. 7.

composition of Washington may be fore-shadowed by the local school enrollment statistics which reveal that in October 1957 three-fourths of all pupils in the Washington public elementary schools were nonwhite.[7] Similarly, in New York City nonwhites had increased by 27 percent from 1950 to 1957. Thus, over that recent period, nonwhites increased by 206,000 while whites decreased by 302,000 even though whites in the aggregate outnumbered nonwhites 9 to 1 in New York City at the beginning of the period. Today there probably are over a million nonwhites living in that city; and it has been predicted that by 1970 the combined nonwhite and Puerto Rican count will have more than doubled the number present in 1950. Over the same period the white population of New York City is expected to decline by 750,000.[8] According to the National Housing Inventory conducted by the United States Bureau of the Census the city of Chicago gained a net of about 83,600 nonwhite households and lost 37,800 white from 1950 through 1956. Likewise, the city of Philadelphia grew by 25,700 nonwhite households as against a loss of 9,800 white. While it is not implied that nonwhites are increasing with corresponding rapidity in all cities across the country, it appears that the influx to cities continues unabated.

Regional Differences. Although marked shifts in the nonwhite population are occurring throughout the nation, the tempo at which the redistribution is proceeding varies from region to region. For example, SMA's in the West are experiencing much the fastest rate of nonwhite

growth. During the past decade non-whites in western SMA's increased by 128 percent. SMA's in the North Central States experienced substantial relative nonwhite increases also, 67 percent, while those in the Northeast grew by 49 percent. Although the South still had by far the largest number of non-whites living in SMA's, those SMA's experienced the lowest rate of nonwhite increase—only 24 percent. In significant contrast the white population in the SMA's of the South increased almost twice as fast as did the nonwhites. This results in part of course from the regional drift of nonwhites away from the South. Basically, these nonwhite regional shifts could materialize only as a consequence of the recent rapid mechanization in southern agriculture and more especially of the widespread employment opportunities arising out of the extraordinary war and postwar industrial expansion. Without the latter, the nonwhites could not have translated the available employment and income opportunities into a realization of their personal ambitions and desires to move to the city.

Search for Economic Opportunity. Why are nonwhites migrating to cities and SMA's? Foremost among the special economic factors producing the steady flow is the desire of nonwhites to avail themselves of the greater employment opportunities open to them in certain localities and regions. This subsumes the opportunity to use and develop their improved education, training, and skill; the freedom to join unions with their promise of equality and a fuller measure of security, without respect to color; the legislative protection of fair employment laws and enforcement machinery in certain cities and states; and the increasing tendency of industries in some areas to employ members of minority groups at all occupational levels on the basis of

[7] A mimeographed table entitled, "Summary of Membership for All School Levels, October 18, 1957" (Washington, D. C.: Public Schools of the District of Columbia, Department of General Research and Statistics.

[8] *Current Population Reports* (Washington. D. C.: United States Bureau of the Census), Series P-28, No. 1073 Rev.; and *Bulletin* (New York City: Department of City Planning), November 22, 1954, p. 2.

their competitive abilities. Sociological factors also play an important role in the nonwhite shift. On the one hand, nonwhites search for a larger measure of freedom from segregation and from related social and economic barriers for themselves and for their children. On the other hand, nonwhites are drawn by the appeal of urban life and its attendant opportunities to that part of the rural population with a pioneering urge to improve its economic and social life.

Economic Contribution. Perhaps it is because so much is said so often concerning the disruptive impact of nonwhite in-migration on our cities that most observers overlook the significant economic and other positive contributions of non-whites. During recent years the economy of the nation and of most cities has expanded tremendously. Although automation and technological progress have increased per capita output substantially the critical need for additional manpower has plagued management the country over. The younger southern Negro has proved to be a huge untapped reservoir which has helped to meet this expanded manpower need. Without his labor to help underpin the fast-growing economy of the city it would have been much more difficult to have reached our present high levels of production. His labors have been employed, moreover, mostly at the drudgery end of the occupational scale. Lacking these urban in-migrants, who would have performed those unenticing but essential jobs?

Although it is prudent to ferret out and identify the factors which produce the myriad social and economic problems and consequences associated with nonwhite in-migration in order that constructive solutions may be reached, it is just as important to recognize the vital part these in-migrants play in our expanding economy. Moreover, it should

be recognized that today's nonwhite urban in-migrant is the counterpart of yesteryear's European immigrant. During the decades up to World War I many of our critical labor needs were met mainly from the vast, unending waves of uneducated and unskilled Europeans to our shores. Historically, each wave of untrained urban newcomers has largely filled the jobs and the housing which others did not want; yet each has been desperately needed in the functioning of the respective economies. It is chiefly the nonwhite, however, who has remained in his present plight of generally poor housing and low income. To an important degree this appears to be the conjoined consequence of the fact that the high visibility of his pigmentation readily sets him apart and that the stigma of "slavery and the early decades of freedom left a people who were generally ignorant and backward, [so that] the color prejudice [which] was the heritage of that slavery permitted a perpetual classification of Negroes and the descendants of the Negroes, whatever their individual characteristics or achievements."[9] If his color were to fade with the second generation as did the white European immigrant's accent, mannerisms, and other distinguishing sociological attributes, it seems likely that a noticeably larger segment of the nonwhite population would have been assimilated in this great melting pot. Indeed, "the barriers of color made assimilation for [nonwhite] Americans who spoke English more difficult than it once was for those whose languages, religions, and cultures were alike alien, but whose skin color was technically white."[10] In this connection it is important to observe that the housing market pattern which char-

[9] Irene B. Taeuber, "Migration, Mobility, and the Assimilation of the Negro," *Population Bulletin*, November 1958, p. 131.
[10] *Ibid.*, p. 132.

acterizes today's urban nonwhite in-migrant also tends to parallel that of the former unskilled European immigrant, which produced nationalistic neighborhoods of Irish, Polish, Jewish, Swedish, Italian, and German settlers. Much as the European immigrant once underwent a process of Americanization the in-migrant nonwhite is now undergoing a process of urbanization. The important difference is that nationalistic neighborhoods tend to change or disappear with time whereas nonwhite concentrations usually persist over succeeding generations.

War's Accelerating Force. Yet, wars often take a tuck in time for society by accelerating social changes much as they hasten medical and technological progress. Thus, just as during and following World War I, a wave of nonwhites moved to the big industrial cities of the North so also did another wave radiate from the South during and after World War II and the Korean Conflict. Those wars facilitated this process in several respects. Perhaps foremost was the urgent need for additional manpower in the labor-shortage industrial cities of the West and North. Young nonwhites from the rural South not only were accepted but also were actively recruited in the home communities by personnel representatives or by relatives and friends employed in distant cities. As a by-product of the personal contact between white and non-white workers on the job, many non-whites were found to be capable workers in those occupational endeavors for which they were qualified so that non-whites generally tended to gain in prestige. Furthermore, many white persons responded to the patriotic appeal that inasmuch as nonwhites shared in the manning of the armed forces they should be accepted more freely in our society.

Transitional Housing. A compelling economic reason for nonwhites' in-migrating to central cities rather than to suburbs is that the bulk of the housing available to nonwhites customarily is supplied by racial transition from white to nonwhite use. In-migrant nonwhites characteristically obtain housing formerly occupied by white families living adjacent to established nonwhite neighborhoods. This practice usually involves concomitantly the filtering-down process, i.e., the procedure by which housing passes along to successively lower income groups as the quality of that housing depreciates over the years. The operation of this housing market phenomenon tends to place those occupying the lowest rungs of the economic ladder in the cheapest housing. Traditionally, nonwhites find themselves economically restricted to housing which has filtered rather far down the scale. Inasmuch as the greatest concentrations of old, spacious, cheap housing abound near the center of metropolitan areas, it is here that most nonwhites perforce find housing accommodations which they can afford. It is important to note that as the transitional process in cities gains momentum it makes available to nonwhites an increasing stock of housing of good or excellent quality.

Recent definitive measurements of the transitional process are afforded by data for nine SMA's from a recent Census Bureau sample survey. Because of the original and important character of that survey, selected data from it have been processed for presentation in the accompanying table, along with certain other comparative data for 1940 and 1950. That survey demonstrates that in the Chicago SMA, for example, a net of about 91,000 dwellings which had been occupied by white households in April 1950 were occupied by nonwhites in

RELATIVE INCREASE IN DWELLING UNITS OCCUPIED BY NONWHITES AS COMPARED WITH ALL RACES 1940-1956 AND NET SOURCES OF CHANGE IN DWELLING UNITS OCCUPIED BY NONWHITE HOUSEHOLDS, APRIL, 1950 TO DECEMBER, 1956: SELECTED STANDARD METROPOLITAN AREAS (SMA's) AND CITIES

Item and source of change	Boston SMA	New York SMA	Philadelphia SMA	Philadelphia City	Detroit SMA	Chicago SMA	Chicago City	Los Angeles SMA	Seattle SMA	Dallas SMA	Atlanta SMA
Occupied units, 1956											
All races	695,800	4,428,000	1,202,600	600,600	1,027,500	1,876,000	1,133,000	2,041,000	278,800	238,800	257,200
Nonwhite	20,500	375,000	155,600	126,400	125,100	259,000	215,000	135,000	10,800	28,100	50,800
Percentage increase in occupied units 1950-1956											
All races	7.7%	17.3%	18.2%	2.7%	24.0%	16.7%	4.2%	41.7%	18.0%	27.6%	38.6%
Nonwhite	41.4%	46.9%	26.2%	25.5%	64.8%	69.4%	63.6%	88.9%	32.5%	21.5%	24.8%
1940-1950											
All races	*	20.0%	22.5%	15.3%	34.4%	21.5%	14.5%	54.1%	39.7%	65.6%	37.0%
Nonwhite	*	52.3%	46.8%	53.7%	91.2%	72.2%	72.3%	106.2%	62.0%	33.6%	9.4%
All races, total occupied new units added 1950-1956	66,200	681,000	219,800	42,900	234,700	289,000	76,000	594,000	48,800	70,300	76,100
Nonwhite new private units as a % of all new occupied units	.2%	.6%	.7%	**	1.4%	3.4%	4.6%	1.9%	.8%	2.7%	14.6%

NET COMPONENTS OF CHANGE FOR DWELLING UNITS OCCUPIED BY NONWHITE HOUSEHOLDS (i.e., increase in nonwhite households)

Item and source of change	Boston SMA	New York SMA	Philadelphia SMA	Philadelphia City	Detroit SMA	Chicago SMA	Chicago City	Los Angeles SMA	Seattle SMA	Dallas SMA	Atlanta SMA
Net increase, 1950-1956	6,000	119,800	32,300	25,700	49,200	106,100	83,600	63,600	2,600	5,000	10,100
Nonwhite units added	7,300	144,800	38,900	31,000	58,000	125,500	100,600	69,100	3,600	6,500	14,600
Former white units[1]	4,300	106,800	29,000	24,400	44,100	91,100	75,600	55,600	3,300	2,800	-300
New construction[2]	1,300	32,000	4,800	2,300	6,700	16,500	9,500	14,000	400	4,100	13,100
Public[3]	(1,175)	(27,800)	(3,173)	(2,365)	(3,530)	(6,710)	(5,991)	(2,952)	....	(2,225)	(1,990)
Private (residual)[4]	(125)	(4,200)	(1,627)	**	(3,170)	(9,790)	(3,509)	(11,048)	(400)	(1,875)	(11,110)
Conversions[5][6]	1,100	6,000	7,400	6,500	6,900	12,500	10,500	....	....	600	1,500
Other[5][6]	600	....	-2,900	-2,200	300	5,000	5,000	500	-100	-1,000	300
Nonwhite units lost	-1,300	....	-6,000	-5,300	-8,800	-19,000	-17,000	-5,500	-1,000	-1,500	-4,500
Demolitions[6]	-1,100	-25,000	-4,200	-2,700	-7,100	-13,000	-12,000	-4,500	-700	-1,800	-3,400
Mergers[6][7]	-200	-25,000	-2,400	-2,600	-1,700	-6,000	-5,000		300	300	-1,100

* Not compiled because of the complexity of the SMA.   ** Not computed because sampling variability produced a negative figure.

[1] The net quantity of racial transition from white to nonwhite occupancy reflecting only the difference between the shift from white to nonwhite and from nonwhite to white occupancy for units unaltered as to physical use over the period. In addition, an unknown number of converted, merged, and "other" units occupied by nonwhites in 1956 had been occupied by whites in 1950.

[2] Inasmuch as these sample figures are relatively small, a fairly large degree of sampling variability may apply to each. Figures obtained from the PHA headquarters and local housing agencies.

[3] Includes both federal and non-federal public housing.

[4] Includes change in use as well as structural alteration. Either type of conversion may bring about an increase or a decrease in number of units.

[5] Net change after subtracting total units removed from the inventory from total units created over the period.

[6] Includes units created over the period 1950 to 1956 from quarters which were not dwelling units in 1950, such as a dwelling unit created from a sleeping room in a rooming house; from a nonresidential space such as a store, garage, or barn; from units moved to site during the period, such as a trailer. Losses occur in a converse manner.

[7] Combination of two or more dwelling units into fewer units through change in use or through structural change. While the merger process always decreases the all-race total, there may be an increase in the nonwhite merger units as a result of a concomitant shift from white to nonwhite occupancy of some of the merger units.

Source: 1956 National Housing Inventory, Components of Change 1950-1956, (Washington, D. C.: United States Bureau of the Census), Vol. I, Parts 2-9, Tables 1, 2, and 3; and Vol. III, Parts 2-9, Table 1; 1950 Census of Housing, Vol. 1, Parts 2-6, General Characteristics, Table 17; and 1940 Census of Housing, Vol. II, Parts 2-7, Table 22.

December 1956. This amounts to over one-third of the aggregate housing inventory occupied by nonwhites in December 1956. Over the same period only about 16,500 new units were occupied by nonwhites in the Chicago SMA. Of these, federal- or nonfederal-aided public housing programs provided some 6,700 units. Therefore, only about three percent of all new private housing built and occupied in the Chicago SMA during these six and three-fourths years was occupied by nonwhites. In this connection public housing provided the bulk of the new construction occupied by nonwhites in several of these SMA's over this period. On the other hand, transitional housing formerly occupied by whites accommodated about 107,000 nonwhite households in the New York SMA, 56,000 in the Los Angeles SMA, 44,000 in the Detroit SMA, and 30,000 in the Philadelphia SMA. In each of the nine SMA's which were surveyed the supply of transitional housing far exceeded that of new housing for nonwhites except in Atlanta and Dallas, the two southern SMA's which were included. New private construction for nonwhites actually exceeded the nonwhite household increase in Atlanta, offsetting part of the nonwhite dwellings which were demolished over this period. This suggests that the nonwhite housing market operates according to a somewhat different pattern in the South from that found elsewhere although in the Dallas SMA the racial transition from white to nonwhite use is unmistakable and sizable.

The transitional process is in part a corollary of racial segregation and usually occurs in a stereotyped manner. Thus, in a city where segregation prevails, inmigrant nonwhites tend to join their friends in an established nonwhite neighborhood. Meanwhile, adjacent white residents often relocate in the outskirts of the city or in the suburbs. Moreover, there appears to be a general but not universal predilection on the part of nonwhites for close-in locations both because of convenient access to church, social, recreational, and other essential facilities welcoming nonwhite participation and also because of proximity to sources of employment and transportation advantages to families without cars. Nonwhites are then to a very real degree self-segregating, much as other racial, social, or economic groups tend to be.

Actually, it is said that housing in the process of transition from white to nonwhite occupancy often is improved and maintained better by the new nonwhite buyers than it had been by the vacating white families, partly because many white owners tend to lose interest and motivation when their neighborhood is in process of, or under threat of nonwhite infiltration. Furthermore, instead of depressing the price of houses, there is research which indicates that the pent-up nonwhite effective demand in the face of the severe shortage of dwellings available to them tends to stabilize or enhance the value of low-and moderately-priced housing in transition, although the transitional process is said to usually depress the market in the higher-priced field because of the thin demand at that level among nonwhites. Indeed, even in the moderately-priced brackets, whites sometimes produce a temporary price collapse by departing en masse and thereby suddenly glut the market. Therefore, whether prices rise, stabilize, or fall during the transitional process apparently is largely a function of the breadth and strength of

nonwhite demand.[11] A further collateral aspect of the process of nonwhite transitional occupancy of older housing left in cities by whites is that it tends not only to accelerate but also to underwrite the suburban movement of families which began several decades ago quite independently of the more recent nonwhite ingress.

Disruptive Impact. The mention of any friction or adverse consequences which may occur during the process of nonwhite population redistribution often is studiously avoided not only by some nonwhite but also by some white elements of the community. Apparently this attitude stems from the mistaken belief that a forthright admission of such nonwhite problems is somehow inimical to the nonwhite interests or perhaps tends to obstruct or impede a wholesome solution of those problems. Actually, in the field of racial relations as in all other phases of life it is axiomatic that the identification of the problems facilitates an intelligent solution of them. In this connection it is abundantly evident that repercussions sometimes do attend nonwhite in-migration, again as was true of past immigration of unskilled and uneducated Europeans.[12]

Insofar as revenues are concerned, nonwhite in-migration often produces a dampening tax base in the city and a broadening base in the suburbs. Costwise, it ushers in heavier welfare and other burdens in the city as against a lighter load in the suburbs. Not only are the cities beset by shrinking real estate, personal, sales, and income taxes as this interchange progresses but also crime, health, and welfare costs rise. It is estimated, for example, that Washington, D. C. would have saved over $6 million in 1954 for all welfare services if nonwhite cost ratios were as low as those for whites and that in January 1955 the monthly per capita public assistance payments in Washington ran 71c as against 11c in the surrounding suburbs.[13] Officials not uncommonly find the financial structure of their cities adversely affected by the process. Obviously, to offset the declining tax potential the tax base must be broadened. This may be accomplished, in part, by levying an income tax at the place of employment in order to tap the salary of the city worker who lives in the suburbs. A counter-measure adopted in Washington, D. C. was the introduction in 1949 of a selective 2 percent con-

[11] Belden Morgan, "Values in Transition Areas: Some New Concepts," *The Review of the Society of Residential Appraisers*, March 1952, pp. 5-10; Luigi M. Laurenti, "Effects of Nonwhite Purchases on Market Prices of Residences," *The Appraisal Journal*, July 1952, pp. 314-329; B. T. McGraw, "Urban Renewal in the Interest of All the People," *Phylon*, Spring, 1958, p. 51; and *They Say that You Say* (Philadelphia, Pennsylvania: American Friends Service Committee), November 1955, p. 9.

[12] The very fragmentary list of reports which follows clearly reveals that significant differences currently prevail as between nonwhites and whites: *Uniform Crime Reports, Annual Bulletin, 1957* (Washington, D. C.: Federal Bureau of Investigation), Table 44; Marcus S. Goldstein, "Longevity and Health Status of Whites and Nonwhites in the United States," *Journal of the National Medical Association*, March 1954, pp. 83-104; *1956 Biennial Report of Families and Children Receiving Aid to Dependent Children in the District of Columbia* (Washington, D. C.: Department of Public Welfare), Tables 1, 2, 5, 8, and 9; two unpublished tables, entitled, "Public Acceptance Recipients, by Race and Category of Assistance, 1956-1957" and "Children Receiving Child Welfare Service by Charge (Delinquent and Dependent)

and Race, 1953, 1957, and 1958" (Washington, D. C.: Department of Public Welfare), R & S No. 101,558; *Vital Statistics Summary, 1956* (Washington, D. C.: Department of Public Health), p. 50; and *Report of the Subcommittee to Investigate Public School Standards and Conditions, and Juvenile Delinquency in the District of Columbia of the Committee on the District of Columbia* (Washington, D. C.: House of Representatives, 1957), Tables 12-16 and pp. 8, 9, 18, 24, and 44-46. Parenthetically, there is said to be a concerted effort in some quarters to eliminate the separate reporting of nonwhite statistics even though this would remove a very valuable tool which is employed extensively in coping with the present plexus of nonwhite problems.

[13] *The Relation of Dependency to Population Composition and Job Opportunity in the District of Columbia* (Washington, D. C.: Department of Public Welfare, April 4, 1956), p. 2; *Metropolitan Washington Assistance Costs* (Washington, D. C.: Department of Public Welfare, May 1955), p. 7; and *District of Columbia Appropriations for 1959, Hearings Before the Subcommittee of the Committee on Appropriations, Second Session on H.R. 12948* (Washington, D. C.: United States Senate, 1958), pp. 48 and 52.

sumer sales tax. This sales tax was extended to additional consumer items in 1954 and again in 1956. Further revenue was produced by initiating a personal income tax in 1947, by increasing the rate in 1954, and then by lowering the personal income tax exemption in 1956.[14] On the real estate front, relatively less new construction takes place in the core cities except for high-rise office buildings, hotels, larger apartment buildings, institutional buildings, and commercial garages so that the real estate tax base can be broadened mostly by progressive raises in the tax rate or by disproportionate valuation hikes in existing properties. Thus, in Washington the real estate taxes were raised by rate increases in 1948, 1950, 1954, and 1957 and by a property reassessment boost effective in 1959.[15] In the suburbs across the land, on the other hand, there were new construction booms in all tax-generating segments— houses, smaller apartments, motels, major shopping centers, stores, small office buildings, industrial buildings, and gasoline stations. Indeed, over three-fifths of the total dollar amount of all new construction in SMA's takes place in their suburbs, although in 1950 only two-fifths of the population resided there.[16] Hence, the influx of large numbers of lower-paid nonwhites to the core city tends to act as a depressant to the financial and social structure of the city. Possible solutions and ramifications of this dilemma are touched on in subsequent paragraphs.

[14] *State of the Nation's Capital: A Report to the Congress* (Washington, D. C.: Board of Commissioners), February 1958, Section 13, pp. 7 and 8.

[15] *Ibid.*, pp. 2 and 3.

[16] Mary F. Carney, "Suburban and Central City Building in Metropolitan Areas, 1957," *Construction Review*, May 1958, pp. 13-16. However, these building permit data tend to understate building activity in the suburbs as compared with their central cities because of the uneven completeness of reporting as between the two types of areas. See also, *1950 Census of Population* (Washington, D. C.: United States Bureau of the Census), Vol. I, Table 27.

Economic Status. Often whites who observe nonwhites living in concentrated areas of squalor and deprivation blame them for these conditions whereas actually these nonwhites are said to be more nearly the victims than the cause of blight.[17] One of the most basic solutions to the unequal housing accommodations as between whites and nonwhites would be afforded by a better equalization of incomes of the two races through unrestricted employment opportunities to both. Thus, should the incomes of nonwhites improve materially, they would be economically able to afford not only better housing, but also more adequate food, clothing, medical care, education, and other essential goods and services, and thereby should be able to avoid many rending elements of their environment which now plague them. Equal employment opportunities, including equal pay for equal work is both a democratic and a sound business principle which has spread markedly in recent years. Adoption of this principle has been spear-headed largely under the leadership of such forces as labor, industry, state and city fair employment practices legislation, and federal policy and executive action. Still, the long-standing pay differential and job restrictions observed as between white men and white women workers suggest that many years will pass before nonwhite earnings and employment opportunities and levels will match those of whites.

In this regard, however, nonwhites have come a long way in narrowing the income gap since 1939. By 1956, for example, the median wage income of nonwhite males had risen an estimated 421 percent as compared with only 283 percent for whites. The nonwhite wage gains resulted partly from the tendency of lower-paid jobs to enjoy the most

[17] Morgan, *op. cit.*, cf. p. 6.

rapid relative increases and partly from the simultaneous progress nonwhites have made in qualifying for and getting better jobs. Yet in 1956 the $3,250 median family income of all urban nonwhites still lagged far behind the $5,413 median for whites. Furthermore, the income of only 24 percent of the nonwhite urban families exceeded $5,000 as compared with 56 percent[18] of the urban white families. It is obvious, therefore, that nonwhites generally are at a real financial disadvantage in competing with white families for housing and other goods and services. Although the income limitations are serious, the predicament of nonwhites often is aggravated by the unwillingness of many owners to rent or sell better housing to nonwhites no matter how fully they meet the test of capacity, credit, and character. In most other markets for essential consumer goods, nonwhites are free to compete on equal terms with whites for goods which they want and can afford.

A further review of comparative income figures for selected cities reveals some marked contrasts favorable to nonwhites. Thus, in 1949 the median income of Negro males actually exceeded that of white males in the Johnstown, Pennsylvania SMA, ($2,446 compared with $2,338) largely because over half of those Negroes were employed in the high-paying steel industry. In the Saginaw, Flint, Racine, and South Bend SMA's, moreover, Negro males earned from 90 to 98 percent as much as did the white males. In these four SMA's the Negro males were employed preponderantly in the automobile, steel, or machinery industries. Median incomes of Negro males ranged upward from $2,300 to $3,119 in the 15 SMA's with the high-

est male Negro earnings.[19] With improvements such as these, nonwhites are financially able to afford relatively adequate housing and other commodities.

Another aspect of the income and economic differentials between whites and nonwhites pertains to the tendency of nonwhites to fill unskilled and lower-level jobs despite their recent advances in education and training. In this connection, nonwhites collectively have begun to climb the occupational ladder at a noticeable pace. Still, there is no demand for cotton pickers in the city. Thus, many unskilled in-migrant Negroes from the agrarian segment of the deep South are so nearly on the fringe of the urban labor force that they find themselves at the bottom of the pile in their search for work. Individually, their vertical mobility comes slowly. Yet this occupational up-grading can and frequently does come. In Washington and in other cities for example, many Negroes have broken through the color barriers which once restricted them to such traditional unskilled jobs as maids, custodial workers, and laborers. Now many of them work alongside their white fellows as stenographers, clerks, salespersons, craftsmen, streetcar operators, technicians, and professionals and are employed throughout the federal government in varied capacities.[20] It would be imprudent to imply, however, that nonwhites have attained occupational equality with whites. Moreover, tradition still significantly affects the employment

[18] *Current Population Reports* (Washington, D. C.: United States Bureau of the Census), Series P-60, No. 27, Tables 2 and 28; and unpublished data from which Table 2 was compiled.

[19] *1950 Census of Population* (Washington, D. C.: United States Bureau of the Census), Vol. II, Part 1, Table 185; and Parts 14, 22, 38, and 49, Table 83.

[20] *Five Years of Progress, 1953-1958* (Washington, D. C.: President's Committee on Government Contracts); *Merit Employment in Chicago* (Chicago, Illinois: Commission on Human Relations, February 1956); *The Washington Post and Times Herald*, February 3, 1957; *Current Population Reports*, Washington, D. C.: United States Bureau of the Census), Series P-50, No. 66, Tables 3 and 4; and John Hope II, "The Employment of Negroes in the United States by Major Occupation and Industry," *The Journal of Negro Education*, Summer 1953, pp. 307-321.

of nonwhites in many occupations and areas.[21]

Under-employment and unemployment directly and adversely affect nonwhite income. Inasmuch as nonwhites often obtain only part-time and part-year employment and frequently below their capacity or skill, their earnings suffer by comparison with whites. Indeed, if unemployment among white workers were to rise from their 5.6 percent level to the 12.7 percent reported for nonwhites in March, 1959,[22] very likely bold, positive, remedial measures would be invoked. Yet full employment of nonwhites probably will defy attainment as long as they are not accorded, or do not respond to unrestricted training and employment opportunities as fully as do whites. Nonetheless, the unprecedented prosperity and economic growth of the last decade and a half probably have done more than anything else to check unemployment among nonwhites. In fact, the present high nonwhite unemployment levels probably are moderate compared with what they would be in a languishing economy. This may be explained in terms of the economic maxim that Negroes usually are the last hired and the first fired.

It is, therefore, no economic accident that the central cities increasingly become the repositories of nonwhites. Actually, the lower income levels of nonwhites may be expected to perpetuate this trend for years to come for the simple

reason that those core cities contain vast stocks of old, spacious, deteriorating, cheap housing which they can afford. This does not mean that all used housing involved in the process of transition to nonwhites is substandard. Some of it is of good or excellent quality. Much of the newer, more expensive housing in the suburban fringes, however, may be expected to remain beyond the financial capacity of most nonwhites until they attain levels of education and training comparable to those of the whites and until and unless their employment patterns parallel those of whites.

Solutions. How can the nonwhite problem of our cities, as well as of our nation, be solved? A growing number of observers feel that one solution to this knotty problem is the acceptance of nonwhites on the basis of their individual economic, social, and moral merits, just as whites are accepted, rather than prejudging them collectively on a racial basis. However, even such an irrepressible thinker and champion of nonwhite rights as Charles Abrams, states that, "integration cannot be achieved overnight; [that] outlawing discrimination [by federal action] in one- or two-family houses would probably be premature in most places and [would] invite difficulties; [and that] tyranny is as unjust when practiced by the minority in the enactment of laws in which the majority does not yet believe as it is when the majority imposes laws by which the minority is to be oppressed."[23]

Moreover, for economic reasons it is obvious that the acceptance of nonwhites on the basis of their individual merits would not mean that nonwhites would be dispersed into white neighborhoods in accordance with their 1-in-9 ratio to the total population. As was pointed out

[21] *Employment in the Hotel Industry* (New York City: New York State Commission Against Discrimination, 1958); *Railroad Employment in New York and New Jersey* (New York City: New York State Commission Against Discrimination, and Newark, New Jersey, New Jersey Department of Education, 1958); Irving Babow and Edward Howden, *A Civil Rights Inventory of San Francisco* (San Francisco, California: Council for Civic Unity of San Francisco, 1958), recapitulated in "Two State Reports on Job Discrimination" and "Minority Worker Hiring and Referral in San Francisco," *Monthly Labor Review*, October 1958, pp. 1125-1136; and Taeuber, *op. cit.*, pp. 143-144.

[22] *Current Population Reports* (Washington, D. C.: United States Bureau of the Census), Series P-57, No. 201, Table 5; and Series P-50, No. 66, Tables 5-10.

[23] Charles Abrams, *Forbidden Neighbors* (New York: Harper & Brothers, 1955), pp. 355, 357, and 385.

earlier, not only are their incomes generally too low to enable them to infiltrate proportionately but also nonwhites usually prefer the practice of self-segregation, at least up to this time. Probably it would require many years—perhaps many generations—before the entire nonwhite population could become up-graded to the economic and social levels of whites. Of course, this solution envisions both equal privileges and equal responsibilities for nonwhites. Because of the inertia-drag of custom it is said that many nonwhites tend to lag noticeably behind whites on both scores.

Various other measures have been adopted in certain areas to bridge the long interim period of adjustment. For example, New York State enacted the Metcalf-Baker Law in 1955, a law which prohibits racial segregation and discrimination in all housing receiving any city, state, or federal subsidy or assistance in the form of land write-down, tax concessions, or mortgage guaranty or insurance. In this connection some type of anti-discriminatory housing legislation has been passed also by at least 12 other states and 30 leading cities.[24] The present New York State law is said to affect 5 percent of the housing in the state.[25] In 1957 New York City also passed the Fair Housing Practices Law which extended the anti-discrimination provision to all new and existing rental housing of three or more units and to the sale of all new one- and two-family houses built in developments of 10 or more.[26]

[24] *Nondiscrimination Statutes, Ordinances, and Resolutions Relating to Public and Private Housing and Urban Renewal Operations* (Washington, D. C.: Housing and Home Finance Agency, October 1958), entire report, and especially pp. 38-46 for New York laws.

[25] Estimate contained in a table entitled, "Proportion of Dwelling Units under Jurisdiction of New York State Commission Against Discrimination, December 31, 1956," (New York City: Commission on Intergroup Relations).

[26] *Fair Play in Housing for Everyone* (New York City: Commission on Intergroup Relations), p. 5, quoting from Local Law 80, New York City, 1957, Chapter 41, Title X, Section X41-1.0, b.(1) of the Administrative Code.

This law may serve as a bellwether for certain other cities.

A broad-gauged housing program is a further important device of which cities may avail themselves. This is needed because of the degrading influence which substandard housing exerts on the daily life of many nonwhites who are confined to such dwellings. To combat these effects any city may employ the better housing tools which are now available. Among these is the Urban Renewal Administration which assists in the transformation of slums and blight-ridden areas into stable, wholesome housing and neighborhoods. The Public Housing Administration provides direct financial aid in housing nonwhites—and whites—who are clearly unable to afford satisfactory quarters without public assistance. Admittedly, however, the PHA is unable to assist all who cannot afford satisfactory housing. The Federal Housing Administration and the Veterans Administration operate in the private housing market by encouraging the provision of good housing in well-planned neighborhoods which resist blight and decadence. Since the Federal Housing Administration makes no direct loans but only insures the repayment of home loans extended by private lending institutions, it is optional with those institutions as to whether they lend to nonwhites.

Furthermore, city governments can avail themselves of various means of effectively coping with problems associated with nonwhites. Some cities have taken such steps as the fusing of the separate efforts of white and nonwhite community leaders into the formation of an area-wide metropolitan committee to deal with common problems; the adoption of a comprehensive and modern system of codes and ordinances relating to building, occupancy, and housing; zoning ordinance revisions; major, long-range street

plans; vigorous code enforcement including the demolition of substandard houses; paving streets or providing sewer and water in substandard neighborhoods lacking them; strengthening of any inadequate administrative organization units; enlisting the support of civic and church organizations in any urban renewal or neighborhood improvement programs; the application for a federal loan and grant to assist in the development and execution of an urban renewal project; and the application for any needed low-rent Public Housing Administration units.

What if measures are not initiated to combat the concentration of nonwhites in the central cities? Inevitably those cities will tend to tap the incomes of the more prosperous residents of their suburbs—often called the city's bedroom. This the central city can do directly by taxing the suburbanite's income at his place of employment. Or it can be done obliquely by seeking federal and state aid in such forms as urban renewal, public housing, school funds, welfare assistance, etc. This aid devolves upon the shoulders of all taxpayers, and more especially upon the suburbanites who enjoy higher income levels. Moreover, taxes on meals, food, cigarettes, liquor, and other items are being imposed in increasingly to meet the rising city costs.

The ballot is still another means by which nonwhites in cities may advance themselves. Although for all practical purposes the bulk of the nonwhites in most parts of the South has been virtually without franchise, a growing segment of the nonwhite population in some southern areas is now exercising this right.[27]

[27] H. D. Price, *The Negro and Southern Politics*, (New York: New York University Press, 1957), pp. 3-6; V. O. Key, Jr., *Politics, Parties, and Pressure Groups* (New York: Thomas Y. Crowell Co., 1958), pp. 651 and 674; Richard M. Scammon, "How Will Negroes Vote?" *The New Republic*, September 16, 1957, p. 14; and Margaret Price, *The Negro Voter in the South* (Atlanta, Georgia: Southern Regional Council, 1957), pp. 1-5.

Outside the South, Negroes have been found to participate freely in elections. For example, New York City, Chicago, Philadelphia, and Detroit have elected Negro Congressmen. New York boasts a Negro President of its Manhattan Borough. There are now several Negro federal judges. Scores of nonwhites are elected members of city councils and state legislatures. Furthermore, in many urban industrial areas of the North and West, as well as a few of the South, the nonwhite vote is being courted increasingly with the Negro population influx. It has even been predicted that by 1970 Chicago may have a Negro mayor.[28]

Summary

The rapid nationwide surge of nonwhites from the rural South and into our big, industrial cities of the North and West stands as one of the most significant socio-economic phenomena of our time. In this massive population redistribution, nonwhites increased relatively more than twice as fast as whites in all standard metropolitan areas and five times as fast as whites in the core cities of those SMA's during the past decade. Data available from scattered, current, special censuses suggest that the city-bound stream of nonwhites which was sparked by World War II and the Korean Conflict continues today at an undiminished tempo. While in-migration supplies the bulk of the nonwhite gains in cities, the rate of natural increase among nonwhites now sharply surpasses that of the whites. SMA's in the West experienced a much faster rate of nonwhite growth that those of other regions although even in the South nonwhites in SMA's increased at a faster rate than they did in the nation at large.

The nonwhite out-pouring arises from a complex of social and economic factors,

[28] "When Negroes Move North," *U. S. News & World Report*, April 13, 1956, p. 29.

foremost among which is the stimulus of new-found employment opportunities previously unavailable to them. While this was born first out of the critical war-time shortages an unprecedented industrial expansion later nurtured and augmented the city demand for nonwhite labor. Simultaneously, a long-felt restlessness dislodges nonwhites from their southern homes as they search for a larger measure of freedom from segregation and from related social and economic barriers for themselves and their children. As the nonwhites are transplanted from their agrarian economy they inject an initial and disruptive impact on our cities and this is widely noted. But on the other hand, they contribute significantly by undergirding the expanding economy of the nation, an economic fact which many observers overlook. These are the workers who usually perform the onerous but essential jobs, low in pay and in social status. They are, moreover, the counterpart of yesteryear's unskilled European immigrant. Unlike the European immigrant, however, the nonwhite in-migrant often fails to be assimilated, largely because of the high visibility and permanence of his pigmentation coupled with an adverse "heritage of that slavery [which] permitted a perpetual classification of the Negroes and the descendants of Negroes, whatever their individual characteristics or achievements."

The nonwhite in-pouring typically involves the process of racial transition whereby punctuated concentrations of nonwhites spill over into adjacent white neighborhoods. As these nonwhite islands fan out, they displace white families who tend to retreat to the outskirts and the suburbs. Inasmuch as the nonwhite influx ordinarily affects older,

cheap housing the filtering -down process likewise is at work. However, it is said that where effective demand is strong, housing in the medium- or low-priced brackets which undergoes racial transition does not usually decline in value although inter-personal conflicts not uncommonly occur as the neighborhood complexion changes. Higher costs of city administration accompany the change, whereas most tax-generating segments of the core-city tend to shrink; yet the tax base broadens in the suburbs in which the higher-paid whites tend to relocate.

The economic status of nonwhites has improved markedly since the outbreak of World War II. Nonwhite incomes have risen dramatically and relatively much more precipitously than those of whites. Furthermore, they have enjoyed substantial occupational up-grading in many cities. Yet noticeable adverse differentials still persist between nonwhites and whites in income levels, in occupational distribution, and in incidence of unemployment. Hence, for many years nonwhites may expect to be restricted economically in obtaining housing and other goods and services.

Possible solutions to the nonwhite dilemma, according to some observers, lie in their hope of being accepted on the basis of their individual merits and capacities just as whites are accepted instead of being pre-judged collectively on a racial basis; in further improvements in their levels of education and training; in limited anti-discriminatory legislation in some areas; in a broad-gauged housing program designed to combat the degrading influences which substandard housing exerts on the daily life of many nonwhites; in several concerted city actions designed to weld together the separate efforts of white and

nonwhite community leaders into the formation of an area-wide metropolitan committee to deal with common problems; in the use of the franchise as a tool in achieving a larger voice in self government; and through other constructive measures, some of which are discussed in the foregoing paragraphs.

Of passing interest to readers of this Journal:

Land Economics was selected, out of an impressive list of over 16,000 titles, to be one of the 250 periodicals chosen and on display at the Moscow Exhibit for the six-weeks' period from July 25th to September 5th, 1959. Copies of the most recent issues occupy a prominent place on an umbrella-shaped kiosk in a court between the first and second buildings on the Exhibition grounds.

The United States Information Agency of the Department of State, in notifying the Editor of this recognition of the magazine, wrote: "American industry, private institutions and the United States Government joined together to put on the Exhibit. The themes represented are: America Learns; America Explores Man and the Universe; America Creates; America Plays; and American Community Life.

Philosophy and Objectives of Watershed Development

By S. V. CIRIACY-WANTRUP*

This article was originally prepared for and presented by the author as the lead-off paper of the Symposium on Economics of Watershed Planning which was held at Knoxville, Tennessee in June 1959. This event was sponsored jointly by the Southeast Land Tenure Research Committee, the Tennessee Valley Authority and the Farm Foundation and was planned for by the Subcommittee on Water Resources of the Southeast Land Tenure Committee of which Professor George S. Tolley, North Carolina State College, was Chairman and guiding spirit. In describing the purpose of the Symposium and the prospects for publication of the papers Professor Tolley writes: "The Symposium was an occasion for representatives of public agencies and universities to take stock of watershed development. It gave a picture of present development and planning procedures and searched for guides to the future. Representative of topics of the twenty main speakers and their discussants were the following: relation of watershed planning to overall resources development; potential extent of watershed development; the watershed as an entity for planning; procedures of interdisciplinary project planning teams; data improvement; original research reports on programming and other mathematical techniques as planning aids; political and legal problems in watershed development; and survey of persons of experience on possibilities for improving watershed development. The Symposium was oriented to small watershed development as exemplified in the Public Law 566 program but river basin and big dam approaches were considered to some extent. The entire proceedings will be published by Iowa State College Press early in 1960."

1. *The Watershed as a Unit in the Social Sciences*

HAVING no claims to the status of a philosopher—being a mere economist—I puzzled a little about the title of my assignment. Rightly or wrongly, I concluded that the planners of this symposium wished the lead-off paper to deal with the essential concepts and principles that underlie the economics of watershed development and that need to be considered for decision-making in public policy. Discussion of these concepts and principles will have the consequence that a number of topics are touched upon which will be treated more thoroughly in subsequent sessions.

A watershed has clear conceptual unity in hydrology, physical geography, and other natural sciences.[1] It is not self-

evident that, as a corollary, a watershed is also a logical unit for understanding and policy making in a social science context. There are many examples in the past and present when large and small rivers and swampy valleys have been boundary lines and barriers to social intercourse rather than arteries of communication and the lines to which settlement gravitates. Frequently, the upstream part of a watershed is occupied by a social group different from the one occupying the downstream part; and political boundaries bisect watersheds at the piedmont zone. In some parts of the world the struggle between lowland people and mountain people has continued for centuries.

As a proposition, I should like to submit that the watershed has emerged rather recently in the social sciences as a unit of understanding and policy making. This emergence appears closely connected with technological change and with shifting demands for the main

* Professor of Agricultural Economics and Agricultural Economist in the Experiment Station, University of California, Berkeley, California, and on the Giannini Foundation.

[1] Sometimes only the dividing ridges are defined as the watershed. This narrow definition is now generally being replaced by a definition that includes the whole area between dividing ridges.

products of a watershed in the course of general economic developcmnt.

The most significant technological change I have in mind was the discovery of large-scale uses for electric energy and of the role of falling water in producing such energy. The second change was the invention—or possibly rediscovery—of concrete and of its reinforcement through steel rods. These two changes made high dams both technologically possible and economically feasible. The experience with the construction of high concrete dams, together with the development of large earth-moving machines powered by gasoline and diesel engines, led to the use of high earth-filled dams where they were cheaper or more suitable—for example, in earthquake areas—than concrete dams. All these technological changes took place during the last quarter of the nineteenth and the first quarter of the twentieth centuries. During the same period and related to these changes, the demand for the main products of a watershed—hydroelectric power, water, timber, livestock, agricultural crops, and recreation—increased greatly. This increase in turn gave new significance to the control of floods, to soil erosion, to sedimentation of reservoirs and canals, to salinity, and to drainage— in other words, to the pervasive problems of water-quality, in contrast to water-quantity management. Water-quality management becomes increasingly important as the demand for products of watersheds increases. This is a special case of the over-all problem of waste disposal when a population of organisms increases in size and density. The atomic age will pose this problem on a gigantic scale. These physical and economic interrelations force considerations of the watershed as a unit in economic understanding and policy.

From this sketch of the emergence of the watershed as a concept in the social sciences, two important conclusions may be drawn. First, the physical and economic interrelations that make the watershed a unit in the social sciences operate largely on the side of production and not on the side of consumption. Consumption of the products of a watershed may take place largely outside of it, and such consumption need not be interrelated. To regard a watershed as a unit in consumption is not required by the physical and economic inter-relations just observed. Policies based on this misconception may hinder rather than facilitate watershed development. Let me give a few illustrations.

Important parts of water law implicitly regard the watershed as a unit in the use and consumption of water. The riparian doctrine, in contrast to the appropriation doctrine, is generally opposed to water export from the watershed in which the water originates. Ground-water laws in many states and the areas and counties of origin laws in California permit water export only as long as it can be regarded as "surplus" water. The California laws require reservation of all potential future water requirements for the watershed.[2] Sometimes it is claimed that the watershed should have the first rights to the hydroelectric energy that is produced within its boundaries.

My second conclusion is that a watershed as a helpful conceptual unit in the social sciences is not immutable. It is a concept of economic dynamics and not of statics. By dynamics I do not mean reference to time merely by dating, but an explicit consideration of changes in technology, preferences, and institutions. In the same way that the concept of a watershed has emerged in the social

[2] For details on these laws, see S. V. Ciriacy-Wantrup, "Some Economic Issues in Water Rights," *Journal of Farm Economics*, December 1955, pp. 875-885.

sciences in the course of technological and general economic change, so its significance may well be affected by similar changes in the future. Again, let me illustrate. A presently foreseeable technological change that may have such results is artificial modification of precipitation. In weather-making, a single watershed, even of large size, may not be an appropriate unit for understanding and policy making. The boundaries of individual watersheds, which may be different for surface and ground water, change over time through geologic forces and drastic action by man.

At this point it may be objected that the self-sufficient irrigation cultures which have existed in many parts of the world during long periods of history prove the watershed concept of early origin, static over time and including inter-relations in consumption as well as production. It may be well, therefore, to note in passing that existing knowledge about the early irrigation cultures tends to confirm my proposition and conclusions. In these cultures the modern watershed concept did not exist and this absence may well be related to some of the difficulties they encountered.

The early irrigation cultures were located in the alluvial plains of major river basins. They relied on periodic flooding and low diversion dams. Levees and water distribution systems were highly developed; but high storage dams, upstream water management, especially quality management, such as silt control, salinity control, and drainage were absent. It is quite possible that the physical difficulties caused by this absence were not the only ones and that the actual destruction of some irrigation cultures were caused by military attack from outside. But such an attack came generally from the less civilized people, the "barbarians" in the upper watershed, attracted by relatively high economic development downstream. In this sense, one may regard such attacks as a consequence of the absence of economic and political unity in the watershed.

2. *Watershed Policy and Watershed Projects*

When tracing the emergence of the watershed as a unit in the social sciences, I referred to public policy rather than public projects. The difference between watershed policy and watershed projects is not merely semantic but highly important conceptually and operationally. The economics of watershed projects—for example, benefit-cost analysis and other quantitative techniques for evaluating such projects—comprises only a segment, and sometimes only a small segment, of the economics of watershed policy. In such project analyses, significant aspects of watershed policy are mentioned, if at all, as "institutional constraints." I should like to submit that variation of such constraints (in other words, whether they are means or obstacles of social welfare) is one of the most important aspects of the economics of watershed development.

In modern Western society, watershed development is accomplished largely through decentralized decision-making of many individual agents, both private and public. The public agents are: for example, flood control, drainage, irrigation and conservation districts; municipalities; and federal and state forests. The private agents range from small subsistence farms to large commercial forest holdings, public utilities, and industrial corporations. These agents are subsectors in organizing and operating Western economies.

The rules of the game under which subsectors make decisions become operational largely through property institutions and the price system. The design,

performance, breakdown, and malfunctioning of these two systems are therefore of special interest in the social sciences. The measurement of performance and the meaning of breakdown and malfunctioning will be discussed presently for the price system (Section 4).

The ground rules and their continuous adjustment are the domain of policy decisions. Individual public watershed projects may be regarded as subsectors in this sense, subject to ground rules not greatly different in principle from those applying to other subsectors.

3. *Objectives and Criteria of Watershed Policy*

In proceeding now to the objectives of watershed policy one encounters what might be called the problem of unity of social objectives and criteria. The objectives of watershed policy cannot, in principle, be divorced from those of other economic policies. Such objectives are interrelated. Social welfare criteria are no different in watershed policy than anywhere else.

Since Pareto,[3] economists have taken a special interest in optimizing social welfare and in the criteria for such optimizing. The present occasion is not suited for discussing the pros and cons of this literature. It must be noted, however, that optima of social welfare and formal criteria for optimizing are constructs in the sense of useful scientific fictions.[4]

[3] Vilfredo Pareto, *Cours d'Economique Politique* (Lausanne, Switzerland: F. Rouge, Libraire-Editeur, 1897).

[4] "A fiction is permissible in science if its character is clearly understood. A fiction is deliberate, conscious deviation from reality. A fiction, however, is not a hypothesis or theory. By itself, a fiction is not intended to be validated by testing with empirical evidence. But a scientific fiction should be useful as a stimulus for or as part of hypotheses and theories which *can* be so tested. That means the test of a scientific fiction is its conceptual usefulness, its expediency, in understanding, explaining, and predicting reality. A fiction becomes mere dogma and, therefore, unscientific if its two characteristics—consciousness of its fictional nature and conceptual usefulness—are obliterated. There are many examples in the history of science of fictions changing into dogma." For more details on this point, see S. V. Ciriacy-Wantrup, "Policy Considerations in Farm Management Research in the Decade Ahead," *Journal of Farm Economics*, December 1956, pp. 1301-1311.

Optimizing is not and cannot be an actual policy objective.

These fictional constructs are conceptually useful as organizing principles for the great number of variables and kinds of relations that must be considered in welfare economics—to decide which ones to bring into the analysis explicitly, which ones to neglect, which ones to combine with others, and which ones to take into account as constraints. Information about variables and relations is insufficient for projecting an optimum expansion path of social welfare over time in a dynamic framework. The actual objective of policy decisions involves successive incremental improvements of the existing state of welfare, considering a limited number of alternatives. The Pareto criterion likewise is suited only for appraising whether an increase of social welfare results, but not for projecting an optimum.

For policy decisions of more limited scope—for example, evaluating individual watershed projects—incremental improvements in social welfare can be determined cardinally. This may be done by comparing hypothetical changes of aggregate national or regional income that can be attributed to alternative projects or parts of them. For policy decisions of broader scope, incremental improvements in social welfare can be appraised only ordinally in terms of directions of changes, the relative speed of changes, and their sequence in time. The issues involved in optimizing versus incrementally improving social welfare will be taken up in more detail later when quantitative analytic techniques are considered as an aid in watershed policy (Section 5).

Having stressed the unity of objectives and criteria in public policies, I must now make a suggestion which, to a few, may appear at first sight inconsistent: In natural resource economics, and par-

ticularly in watershed development, there are some significant conditions which induce characteristic divergencies in the actual welfare performance of the economy from fictional welfare optima. This situation makes it permissible to focus on these significant conditions and their changes when considering policy objectives. Some of these conditions are not less important for other than watershed policies. But most of them may be called special conditions because they affect natural resources and watershed development to a greater degree than other aspects of the economy.

We may ask then what are these significant conditions. They can be systematized in several ways. In conservation economics[5] for example, a differentiation has been made between various classes of benefits and costs and an explanation offered why their allocation to decision-making agents and the incidence among members of a social group lead to characteristic divergencies from a social welfare optimum.

In this paper I should like to systematize these conditions in a somewhat different way because my assignment emphasizes the relation of watershed development to the market economy and the price system. What I should like to do is to differentiate between types of breakdown or malfunctioning of the price system. It is frequently argued that market prices are the signaling system that steers Western economies toward the social-welfare optimum. It is important, therefore, to know in what respect and for what reasons the price system does not and cannot fulfill this function. Under these conditions it becomes necessary for public policy to find substitute or supplementary or countermanding systems.

4. *Watershed Policy and the Price System*

In systematizing types of breakdown and malfunctioning of the price system three major types may be mentioned. All of these are of significance for public policy in general. In degree, the first two are of special significance in watershed policy. They are: (1) Price signals do not exist. (2) Price signals are not received by the agent who makes decisions but are received by others. (3) Price signals are "distorted" in a defined sense.

Price signals do not exist for that part of benefits and costs of watershed development which I like to call "extramarket." Some of them are collective benefits in the sense that they are not divisible in consumption. The scenic values of a watershed unmarred by soil erosion, destructive logging, billboards, and slums are an example. Collective costs are the damage by floods and its risk to the general economy of a watershed rather than to individual properties; and threats to public health, such as malaria, related to drainage conditions.

Some benefits of watershed development, on the other hand, are divisible in consumption. In other words, a price could be charged for their enjoyment by individuals. Many of these goods, however, are free or nearly free institutionally. Recreational facilities offered by public reservoirs and public hunting and fishing are examples of extramarket benefits that are divisible in consumption but are public goods institutionally.

For some time resource economists have been emphasizing the importance of extramarket benefits and costs without reference to particular stages of economic development. Recently John Kenneth Galbraith has called special attention to them with reference to affluent societies.[6]

[5] S. V. Ciriacy-Wantrup, *Resource Conservation: Economics and Policies* (Berkeley, California: University of California Press, 1952), especially Chapters 16-18.

[6] J. K. Galbraith, *The Affluent Society* (Boston, Massachusetts: Houghton Mifflin & Co., 1958).

Great emphasis is given the increasing lack of what is called "social balance" between the products supplied by the market economy and products such as education, defense, parks, and playgrounds which are publicly supplied and financed by general taxation rather than sale. Galbraith reasons that the supply of extramarket goods has an inherent tendency to lag behind the supply of market goods because modern advertising and emulation, which are largely responsible for demand shifts in affluent societies, operate exclusively in favor of market goods.

While in substantial agreement on these points, I should like to note that the problem of social balance between market and extramarket goods is not confined to affluent societies. Furthermore, there are many examples of societies—not excluding contemporary ones— for which social balance might be regarded as threatened by the preponderance of publicly supplied goods.

The historical fact of a great many past and present ratios between the supply of market and extramarket goods raises the question of what is the criterion of an optimum social balance. Galbraith rejects the traditional criterion, namely, that the utility from a marginal increment of productive services devoted to the production of extramarket goods should be equal to the utility of the same increment devoted to the production of market goods. This criterion is rejected because the utility of market goods is what is called "synthesized," whereas utility from extramarket goods is not. A precise optimum in the social balance is regarded as unimportant. The direction in which policy should move to correct this condition is regarded as plain and the distance to be traversed considerable.

It has already been suggested (Section 3) that direction is frequently an acceptable criterion for incremental improvements in social welfare through policy decisions of broad scope. For individual watershed projects one may well consider going further. Whether the economist likes it or not, evaluation of extramarket benefits and costs—and also dismissal of such evaluation—is already a part of the political process. Reports of fish and game departments and other public agencies illustrate these attempts at evaluation. One may have professional doubt about some of the procedures used. Still, the economist may well take an interest in them in order to develop better substitutes. Otherwise, the arguments of well-organized groups interested in market values alone, who dismiss extramarket benefits and costs as intangible, might receive disproportionate attention in policy decisions.

This is not to suggest that *all* benefits of recreational resources could be evaluated. In connection with many such resources, however, market values can be used indirectly—for example, through analyzing data on fees, leases, and real estate transactions. In other cases, measurement in terms of physical units of use, for example, man-days, can be accomplished fairly easily. Values of additional units of use can be approached through questionnaires and the study of behavior in other experimental choice situations. Even such crude and partial measurement is more useful than disregarding these values altogether, or substituting for them some figure based on the expenditures of users for transportation, room and board, guns, fishing tackle, and similar items.

Proceeding now to the second major type of breakdown or malfunction of the price system, what is meant by "price signals are not received by the decision-

making agent but by others?" In classical and neoclassical economics, these problems appear as external economies and diseconomies. In watershed economics, these externalities are discussed largely under the labels "offsite" and "indirect" benefits and costs.

External economies and diseconomies are of many kinds. They may be market or extramarket, pecuniary or nonpecuniary, static or dynamic, reversible or irreversible. In spite of the early article by Ellis and Fellner[7] there is still much confusion between externalities that are merely transfer items in an international, national, regional, or local framework and those that are not. By some authors the terms "pecuniary" and "nonpecuniary" are employed in order to differentiate between transfer and nontransfer items. This is confusing because externalities may be either market or extramarket benefits and costs. If there are multiple decision levels in the same firm, price signals may affect these levels differently. In farming, for example, the tenant may be affected but not the landlord, or vice versa. In such cases the term "externality" becomes, strictly speaking, inapplicable although the breakdown of the price system is of the same type as in other kinds of externality. For public policy it is necessary to ascertain in each particular case what kind of externality is involved. The term itself is of little help for understanding and policy making.

The origin and the incidence of offsite benefits and costs can, to a considerable extent, be influenced in watershed policy through property institutions—especially resources law and taxation. Taxation is used here in a broad sense as including negative taxation, that is, tax bonuses. Such bonuses may consist of depletion or depreciation allowances, reduction of taxes, or outright supports and subsidies. Tax bonuses may be made dependent on fulfillment by the taxpayer of certain requirements regarding his use of resources.

Public districts can be employed effectively to make the influence of resources law and taxation operational. This is a vital area for cooperative research between economics and other social sciences, especially law and public administration. When supplemental and contermanding systems were mentioned previously (Section 3), I had in mind especially resources law and taxation. It is difficult to see, for example, how the relations between upstream and downstream interests and upland and bottomland landowners could be adequately taken into account by the price system.

The problems of indirect benefits and costs in watershed policy have been analyzed elsewhere.[8] It may be mentioned, however, that transfer items deserve careful consideration for several purposes of benefit-cost analysis. They deserve such consideration for purposes of project repayment, interpreted broadly as including cost sharing and financing—in whatever framework, international, national, regional, or local—and also for purposes of project evaluation, provided the transfer is "out of" the framework considered.

Turning now to distortions of the price system, the first difficulty is, of course, to define "distortions." Although most economists may well agree that for the following four basic reasons distortion exists, few would deny that a value judgment is involved in all these cases and that some qualification is needed with respect to degree of distortion.

[7] H. S. Ellis and William Fellner, "External Economies and Diseconomies," *The American Economic Review*, September 1943.

[8] S. V. Ciriacy-Wantrup, "Cost Allocation in Relation to Western Water Policies," *Journal of Farm Economics*, February 1954, pp. 108-129; and S. V. Ciriacy-Wantrup, "Benefit-Cost Analysis and Public Resource Development," *Journal of Farm Economics*, November 1955, pp, 676-689.

The first case of distortion was indicated earlier when the synthetic nature of market demand in affluent societies was mentioned. This is the broad and difficult problem of the interpretation of and divergencies from consumer sovereignty as one of the basic assumptions in any attempt to employ market prices in an economic analysis of alternative states of social welfare.

The second case of distortion is more often explicitly considered in welfare economics than the first. It occurs when the income distribution that generated the market prices employed in welfare statements diverges from an income distribution regarded as the "ideal." The ideal distribution may be an equalitarian one, or it may be some other.

The third case of distortion occurs in the absence of the necessary conditions regarding market organization which must be fulfilled if the price system is to steer the economy toward a welfare optimum. The effects of monopoly, duopoly, oligopoly, and other divergencies from pure competition have been discussed by economists over the last several decades.

Last and not least, the price system may be called distorted if social institutions have lost the identity of "concept" and "structure," to use Sumner's terms.[9] Social institutions affect market prices not merely through income distribution and market organization but in many other ways, both on the side of demand and on that of supply. It was noted previously (Section 2) that social institutions are brought into welfare economics as constraints. Usually, no attempt is made to ascertain whether they have lost identity of concept and structure.

Watershed policy is concerned with all these distortions but no more so than

other public policies. Watershed policy can work in the direction of correcting for these distortions. On the other hand, existence of these distortions suggests caution when market prices are used in quantitative economic analysis.

These points lead me to the last part of my assignment, namely, to appraise the merits and limitations of benefit-cost analysis for evaluating public investment in watershed-development projects.

5. Benefit-Cost Analysis and Watershed Policy

The literature on benefit-cost analysis and related techniques has increased recently by leaps and bounds. A few years ago the professional contributions in this field were a few articles. Now contributions are counted in books. Last year alone three such books were published.[10] Several others are in preparation. This literature is widely quoted and a flattering review article has just appeared.[11] A careful and critical stocktaking will soon become necessary in order to ascertain where we stand and in what direction we should push forward. This is not the occasion for a detailed critique but a few comments are called for by my assignment.

The claim of the "now" benefit-cost analysis is that evaluation of public investment in water-resources projects is viewed for the first time as a problem of optimizing social welfare under budgetary constraints. Much space is devoted, therefore, to formulating criteria for optimizing and to stressing the in-

[9] William G. Sumner and Albert G. Keller, *The Science of Society*, 4 Vols. (New Haven, Connecticut: Yale University Press, 1927).

[10] O. Eckstein, *Water-Resource Development: the Economics of Project Evaluation* (Cambridge, Massachusetts: Harvard University Press, 1958); John V. Krutilla and Otto Eckstein, *Multiple Purpose River Development; Studies in Applied Economic Analysis* (Baltimore, Maryland: John Hopkins Press, 1958); Roland M. McKean, *Efficiency in Government Through Systems Analysis with Emphasis on Water Resources Development* (New York, New York: John Wiley & Sons, 1958).

[11] Julius Margolis, "The Economic Evaluation of Federal Water Resources Development; A Review Article," *American Economic Review*, March 1959, pp. 96-111.

sufficiency of traditional benefit-cost ratios as such a criterion. There is little awareness in this literature that economists in governments and universities who have worked critically with benefit-cost analysis over many years have frequently pointed out the insufficiency of benefit-cost ratios as optimizing criterion. This insufficiency is rather obvious. Why, then, it may be asked, was such use of these ratios not opposed more vigorously? There are several reasons.

First, to those familiar with the operational aspects of benefit-cost analysis, it would seem rather naive to identify optimizing of social welfare (as a useful fictional construct in the sense explained above, Section 3), with actual policy objectives. In time and uncertainty economics, such objectives must be formulated, as we know, in terms of incremental improvements in social welfare. The "new" benefit-cost analysis gives much attention to the numerical value of interest rates and uncertainty allowance that should be used in optimizing. But the real challenge of time and uncertainty economics—namely, how changes of technology, preferences, and institutions are to be taken into account through formulation of objectives and identification of variables and constraints—receives little consideration.

Second, even if optimizing of social welfare is assumed an operational policy objective, the shortcomings of the traditional ratios as criterion for this purpose may well be regarded as of the second order of significance when considering other shortcomings of benefit-cost analysis. These other shortcomings are connected with the reliance on market prices in evaluating benefit and cost streams. The "new" benefit-cost analysis is little concerned with these other shortcomings. In other words, there is no progress in the treatment of extramarket

values, offsite and indirect benefits and costs, consumer sovereignty, income distribution, market organization, and institutional influences. More is needed here than translating economics into the jargon of operations research or renaming externalities as "spill-over effects" or substituting a multiplicity of opportunity cost rates of interest for a multiplicity of market rates.

Third, the particular budgetary constraint that is emphasized by the "new" benefit-cost analysis—namely, that a water-resources budget must be assumed as "given"—is by no means the only one that can be selected. This constraint is merely similar to the one imposed by the present federal administration. Theoretically, optimizing of social welfare for the budget as a whole—not item by item —and relation of investment expenditure to other expenditure and taxation is desirable. Operationally, especially in terms of political decision-making in the legislative branches of federal and state governments, it would seem more relevant to assume that water-resources appropriations are actually made project by project, and that the total water-resources budget is to some extent dependent on the size and number of projects that can muster the necessary political support in the legislatures. Under this assumption, benefit-cost analysis has the more modest but still, highly important functions of ruling out or at least stigmatizing, projects that do not make an incremental contribution to social welfare and of selecting for each project not necessarily the best alternative but at least one of the better ones. For these purposes, ratios can serve.

Evaluation of public investment in water-resources projects or, for short, project selection, is only one of three important problem areas in which benefit-cost analysis may be useful in water-

shed policy. The other two areas comprise the broad problems of repayment of project costs, including problems of cost sharing and financing, and the related problems of pricing those products of a project that are sold (Section 4). The "new" benefit-cost analysis pays little attention to the problems of repayment and pricing. This is in accordance with the emphasis on optimizing and optimizing criteria. One may submit that the contribution of economics as an operationally significant policy science is potentially greater in the areas of repayment and pricing than in that of evaluation. My reason for this proposition is that problems of repayment and pricing remain important after a project has been selected and constructed. This means that benefit-cost analysis can operate in the areas of repayment and pricing with a more complete and better identified set of conditions that can be assumed as "given."

Present trends are toward substitution of linear programming for benefit-cost analysis. The programming techniques presently in use are static even though dating may be employed. The basic mathematics of dynamic programming have been known for some time. They have been used in the conceptual clarification of conservation economics.[12] As yet, no operational applications of dynamic programming are available in watershed economics. It will be interesting to see, when actual results become available, whether the particular advantages of the technique are sufficient to overcome its particular limitations.

Benefit-cost analysis is essentially an informal but flexible programming technique. The informal technique leaves

considerable latitude to the user in exercising judgment, professional competence, and integrity—or absence of these qualities—in selecting and stating his assumptions. In formal programming the assumptions are largely "built in" and concealed from those who are not familiar with the techniques—such as most legislators. Formal programming is superior to benefit-cost analysis in the sense that it can determine cheaply and precisely an *optimum optimorum*. As already implied, the significance of this superiority in time and uncertainty economics is at least doubtful. The numerical precision in determining optima may actually become harmful if it induces greater confidence in them. Calculation must be projected for 15 to 20 years in the future. Such a gestation period is unavoidable between the planning stage and operation. The subsequent pay-out period cannot well be set at less than 40 years. Programming for subsectors in the above sense (Section 2) necessitates great detail in "activities" and "processes." While the conceptual defects of optimizing for subsectors are less than for broader policy decisions, the problems posed by the availability of data for smaller statistical aggregates are usually greater.

Turning now to policies of broader scope, we noted that they can often be appraised only ordinally and in terms of direction of changes, the relative speed of changes, and their sequence in time. Before concluding, one aspect of such an appraisal may be mentioned because it is of particular interest to watershed policy.

It is important for public policy to know whether an observed or planned change sets in motion corrective counterchanges tending over time toward a balance of the initial change, or whether a change sets in motion other changes that are circular and cumulative and

[12] S. V. Ciriacy-Wantrup, "Private Enterprise and Conservation," *Journal of Farm Economics*, February 1942, pp. 75-96.

tend to reinforce the initial change.[13] There is, for example, evidence that in the United States during the last 50 years relations between prices of land, prices of agricultural inputs other than land, and technological change in agriculture can be regarded as a corrective system.[14] On the other hand, Myrdal's well known thesis about the nature of economic change is an example of a circular and cumulative system.[15] There is also evidence that the relation between soil erosion, income, population, and technology is frequently such a system.[16] Such a system, if operating in the upstream portion of watersheds, has physical and economic consequences downstream which have circular and cumulative effects upon the whole watershed.

Watershed policy may well make allowance for the risks involved in this circular and cumulative system. Such allowance cannot be made by adding a few percent to the interest rate used in benefit-cost analysis. Rather, the approach is that of establishing minimum standards in resource use through a variety of institutional approaches. In this way, allowance for uncertainty is built into the formulation of policy objectives itself.

This is merely an illustration for the generalization suggested above (Section 3), namely, that objectives of policy of broader scope may well focus on significant conditions that cause characteristic divergencies in the actual welfare performance of the economy from fictional welfare optima. For the pursuit of such objectives, benefit-cost analysis and relative quantitative techniques must be supplemented by a type of analysis that takes cognizance of research in economic history, in the sociology of value systems, and in the change of social institutions, especially the law. This type of analysis relies heavily on theory but not on economic theory alone; it focuses on time and uncertainty economics but not necessarily through increasing the number of variables and equations in optimizing.

[13] The common terms "equilibrium" and "disequilibrium" are not too well suited for describing these two systems. A corrective system has at best only a tendency toward equilibrium. A cumulative system, under some conditions, is more likely to realize and maintain equilibrium.

[14] S. V. Ciriacy-Wantrup, "Conceptual Problems in Projecting the Demand for Land and Water." Presented before the Land Economics Institute, University of Illinois, June 24, 1958. (In press.)

[15] Gunnar Myrdal, *Rich Lands and Poor, World Perspectives*, vol. 16 (New York, New York: Harper and Brothers, 1957).

[16] S. V. Ciriacy-Wantrup, "Resource Conservation and Economic, Stability," *The Quarterly Journal of Economics*, May 1946, pp. 412-452.

Atomic Power: Research Costs and Social Returns

By EDWARD F. RENSHAW*

IT IS not without hesitation that one attempts an analysis which even at its inception may be out of date. Research, by its nature, is concerned with the isolation of relationships and parameters that are unknown; it is difficult to imagine a situation in which resources are committed that entails greater uncertainty. Uncertainty as to the magnitude of the costs and returns associated with the development of new techniques, however, is not an excuse for ignoring information and expectations which do exist or can be made readily available for guiding decision making; we still want to be as confident as possible that resources invested in research are put to their highest alternative use.

To the extent that an economic value can be assigned to the product of research, the conceptual framework for estimating costs and returns and for choosing between projects is straight forward; indeed, it is not unequivalent to the problems of estimation and choice involved in making other kinds of resource commitments. The basic difference, if in fact one exists, is that management must be even more prepared to revise cost-benefit estimates and alter decisions once additional information becomes available.

A central purpose of this paper is to extend the idea that it is possible to estimate not only the ex post returns from public and private investment in research but that it is also possible to make ex ante estimates of the returns from research not yet completed.[1] The field of nuclear fission power has been chosen for its topicalness and because there exists an abundance of economic and technical projections which can be used to facilitate analysis and minimize the discretion of the economist.

Nature and Magnitude of Costs Incurred in Development of Nuclear Fission Techniques for Production of Civil Power

The development of civil atomic power is one phase of a predominantly military atomic energy program which from its inception in 1940 to the end of the fiscal year 1956 resulted in public appropriations of $15.2 billion,[2] at least several billion dollars of which would have been necessitated by a single-purpose civil power effort.[3]

In order to stimulate a civil power development program the Atomic Energy Commission has undertaken three major programs of financial assistance. First, it has been conducting research and development on promising reactor concepts. Generally, this program has been limited to experimental pilot-plant models. As of fiscal 1957 it has been estimated that between 487.9 and 595.8 million dollars were spent or committed to civil power development under this program.[4] It should be noted that these figures underestimate the true cost of the AEC's experimental pilot-plant program, originally scheduled for completion within five years, since not all of the projects have been completed and the cost estimates for the five original reactors (there are now twelve included in the program)

---

* Department of Economics, The University of Chicago.
[1] For a discussion of the economic principles involved in estimating the social return from public and private investment in research see, Zvi Griliches, "Research Costs and Social Returns: Hybrid Corn and Related Innovations," *The Journal of Political Economy*, October 1958, 419-431.

[2] Atomic Energy Commission, 21st Semiannual Report, January 1957, p. 379.

[3] Richard A. Tybout, "Atomic Power and Energy Resource Planning," *Federal Expenditure Policy for Economic Growth and Stability* (Washington: Government Printing Office, November 5, 1957), p. 771 .

[4] Tybout, *op. cit.*, p. 775.

have increased from 36 to 89 percent over original estimates.[5] Estimates are not available as to the magnitude of research and development expenditures made by private industry in the hope of developing a market for nuclear reactors.

As a part of its second program the Commission covers part of the cost of full-scale nuclear power stations authorized for private operation under the power reactor demonstration program. A total of 139.4 million dollars has been appropriated for this program through fiscal 1958.[6] This figure, however, understates the true cost of the program because of a tax subsidy and shifting of uneconomic power costs on to the consumer. Rapid tax amortization is available for civil atomic power plants though not for other electric utilities. At the state level, established regulatory practice has been interpreted to permit atomic power station operators to charge the higher costs of atomic electricity to consumers via higher rates than would have resulted had conventional facilities been installed. Another device used to shift costs on to the consumer is the charge of nuclear construction costs to research and development, which are commonly considered operating expenses rather than investment.[7]

By way of a third program, the AEC had adopted a method of subsidizing the operation of nuclear plants by announcing a sliding scale of price supports for the purchase of regenerated fissionable materials produced in private atomic power plants. It has been estimated that the buy-back price announced by the AEC May 18, 1957, constitutes an operating subsidy that will provide a revenue of 2.5 to 3.75 mills per kilowatt-hour for reactor operators.

In addition to the three programs previously mentioned, the civil atomic power industry is indemnified by law against public claims of damages from nuclear accidents over certain levels.

While a complete enumeration and identification of the public and private costs involved in the development of economic nuclear power would be both helpful and desirable, such a formidable task is beyond the scope of this paper. For the sake of a simplified analysis, it will be assumed that there are two kinds of costs involved—the cost of research inputs and an opportunity cost incurred as a result of producing uneconomic power during the developmental period. Presumably, both research inputs and operational plants or models are necessary factors in the achievement of economic nuclear power though to some extent they should be conceived as substitutes for each other.[8]

A Benefit-Cost Analysis

Classifying the costs of obtaining advances in nuclear technology in terms of research costs and the cost of uneconomic power greatly simplifies the analysis which is carried out in Table II and its accompanying footnotes.

In Table I can be found various estimates of the projected cost of nuclear power for the years 1955 to 1980. The projected cost decline (col. 1) which is used in the benefit-cost analysis tends to be on the optimistic side. No particular brief will be made for its acceptance. It is merely meant to be illustrative of existing estimates: its primary function is to show how projections, which are made

[8] The analysis herein ignores the problem of combining the two inputs such that the total cost of achieving economic nuclear power is minimized; it also overlooks what might amount to formidable conceptual and practical difficulties in distinguishing between the two kinds of costs, theoretically or in an accounting sense. No brief will be made for having established either a desirable or an important distinction between costs; for purposes of this analysis, the distinction is merely useful.

[5] *Loc. cit.*
[6] *Ibid.*, p. 776.
[7] *Ibid.*, p. 789.

TABLE I—PROJECTIONS OF NUCLEAR POWER COSTS IN THE UNITED STATES: 1955-1980
(*Shown in Mills per KWH*)

Year	Projection Assumed in Table II	Projection Assumed by Karl Mayer[1]		Projection of Davis & Roddis[2]	
		Small plants (less than 50 MW)	Large Plants (50 MW and Larger)	Cost Range	
				Low Cost Estimate	High Cost Estimate
	Col. 1	Col. 2	Col. 3	Col. 4	Col. 5
1955.............	20.0	....	..	20.0	....
58.............	18.6	..	..	....	....
57.............	17.2	..	..	....	....
58.............	15.8	..	..	....	....
59.............	14.4	..	..	....	....
1960.............	13.0	18	13	12.0	50.0
61.............	12.6	..	..	....	....
62.............	12.2	..	..	....	....
63.............	11.8	..	..	....	....
64.............	11.4	..	..	....	....
1965.............	11.0	16	11	9.5	12.5
66.............	10.6	..	..	....	....
67.............	10.2	..	..	....	....
68.............	9.8	..	..	....	....
69.............	9.4	..	..	....	....
1970.............	9.0	14	9	8.0	10.0
71.............	8.6	..	..	....	....
72.............	8.2	..	..	....	....
73.............	7.8	..	..	....	....
74.............	7.4	..	..	....	....
1975.............	7.0	12	7	6.8	8.2
76.............	6.8	..	..	....	....
77.............	6.6	..	..	....	....
78.............	6.4	..	..	....	....
79.............	6.2	..	..	....	....
1980.............	6.0	10	6	5.8	7.0

[1] Karl M. Mayer, "Nuclear Power and the World Market," *Prospects for Economic Nuclear Power* (New York: National Industrial Conference Board, Inc., 1957), p. 48.

[2] Data obtained by graphic approximation. See, W. Kenneth Davis and Louis H. Roddis, "A Projection of Nuclear Power Costs," *ibid.*, p. 23.

anyway, can be used to answer economic questions.

Data on the cost of producing power by conventional means has been computed and projected into the future by Karl Mayer. These projections form the foundation for an economic analysis of the negative and positive savings which are likely to be associated with the pro- duction of nuclear power. Unit savings (Table II, cols. 2 and 4) are obtained by subtracting the projected cost of nuclear energy from the projected cost of conventional energy. The Davis and Roddis estimate of nuclear output (col. 1) is used to obtain estimates of aggregate savings (cols. 3 and 5) which are compounded or discounted by a five percent

TABLE II—PROJECTED GROWTH OF THE NUCLEAR POWER INDUSTRY IN THE UNITED STATES, COSTS, RETURNS, AND RELATED DATA, 1955-80: BASED ON THE PROJECTIONS OF DAVIS AND RODDIS

YEAR	Assumed Growth of Nuclear Power[1] (Billions of Kwh generated)	Assumed Private and Social Cost of Producing Uneconomic Nuclear Power		Assumed Savings from the Production of Economic Nuclear Power	
		(Mills/Kwh)[2]	(Millions of 1967 Dollars)[3]	(Mills/Kwh)[4]	(Millions of 1967 Dollars)[5]
	Col. 1	Col. 2	Col. 3	Col. 4	Col. 5
1955.............	0.0	....	....	....	....
56.............	0.0	....	....	....	....
57.............	0.5	10.2	115.3	....	....
58.............	0.5	8.8	....	....	....
59.............	0.8	7.4	47.2	....	....
1960.............	5.3	6.0	486.1	....	....
61.............	5.9	5.6	64.9	....	....
62.............	8.5	5.2	228.7	....	....
63.............	11.8	4.8	250.3	....	....
64.............	16.4	4.4	306.6	....	....
1965.............	26.3	4.0	568.4	....	....
66.............	37.4	3.6	552.5	....	....
67.............	49.3	3.2	495.4	....	....
68.............	69.6	....	....	1.07	70.8
69.............	102.5	....	....	1.11	102.8
1970.............	148.5	....	....	1.15	147.0
71.............	207.6	....	....	1.22	208.4
72.............	281.2	....	....	1.29	285.1
73.............	367.9	....	....	1.37	375.6
74.............	468.4	....	....	1.44	480.1
1975.............	584.1	....	....	1.52	599.3
76	719.4	....	....	1.61	748.0
77.............	876.4	....	....	1.71	920.1
78.	1,057.8	....	....	1.81	1,118.2
79.............	1,262.8	....	....	1.90	1,339.5
1980.............	1,492.7	....	....	2.00	31,696.1*
Sub Totals.....................................			3,115.4	....	38,091.0
Federal Expenditures for Research and Development in 1967 Dollars[6]....................			2,899.2	....	....
TOTALS.......................................			6,014.6	....	38,091.0

Ratio of Benefits to Cost: 6.33 to 1.00

[1] Davis and Roddis' estimates of installed nuclear capacity converted to billions of kwhrs generated by assuming a 75% load factor. See, Davis and Roddis, *op. cit.*, p. 25.

[2] Assumes that the alternative cost of producing comparable electricity by other means is seven mills per kwhr.This figure represents Mayer's median projected cost of generating power with new steam plants in 1965. It falls between Davis and Roddis' estimate of the present costs of electricity from modern coal-fired plants which are as low as 4½ mills in especially low-cost fuel zones and range up to 9 mills in other areas. A median estimate is believed to be relevant since it is unlikely that any save the larger and more efficient utilities will feel they can afford to bear the cost of producing uneconomic nuclear

rate to 1967 values for comparative purposes.

At this date negative savings (col. 3) are added to a conservative estimate of research expenditures which has been projected separately. The resulting sums are then displayed in the form of a ratio of benefits to costs, not because there exists a special justification for computing such a ratio, but because the ratio has come into vogue as a statistic for comparing investment alternatives sponsored by the federal government.

The preceding sketch tends to gloss over and obscure many of the crucial assumptions which are involved in benefit-cost analysis. Soon after I began exploring this field it became apparent that the returns from the development of economical nuclear fission power are not likely to be as spectacular as many people believe them to be. In order to get this point across in as convincing a manner as possible I have deliberately selected, wherever possible, assumptions which are favorable to the development of nuclear power.[9] (By way of footnote it

[9] A bias of this sort can be rationalized or "justified" on a number of grounds. The benefit estimation procedures of most government agencies are such that benefit-cost ratios in general are biased upwards. Edward F. Renshaw, *Toward Responsible Government, An Economic Appraisal of Federal Investment in Water Resource Programs* (Chicago, Illinois: Idyia Press, 1957). There may be legitimate considerations other than the direct economic returns from nuclear power which would make it desirable for us to invest in this area, such as complementarity between reactor technology for civil uses and for defense, the prestige of having harnessed the atom, and the development of a cheaper source of power for our allies. Mayer's study indicates that nuclear power abroad, particularly in such countries as Italy, France, England, and Germany, will be of earlier and of relatively greater economic significance. On another plane, a general bias will serve to compensate for hidden biases of the analyst; on practical grounds, it sanctions rough approximations to the theoretical ideal and permits less exacting, hence less costly, calculations.

(Continued from page 225)

power on the chance that the experience gained in operating nuclear facilities will permit savings in some future period. The data in column 2 are calculated by subtracting seven mills from the projected cost of nuclear power assumed in this paper, Table I, column 1.

[3] Assumes that stabilized costs (column 2) as a function of the incremental growth in nuclear generation implicit in column 1, are incurred for a period of twenty years beginning in the year generating facilities go into operation. The assumed twenty years in which uneconomic costs are incurred is arbitrarily low; our knowledge of reactor technology is too limited to state with great precision the appropriate depreciation and obsolescence rate. In most nuclear cost studies the amortization and replacement component of capital costs reflects a 25-35 year plant life on a sinking fund basis.

A five percent discount rate is used throughout the analysis to convert costs and returns to 1967 dollars. This rate has the advantage of permitting direct comparisons to be made with investment in agricultural research. With respect to the utility industry the five percent rate is of course slightly higher than the over-all cost of capital to private utilities.

[4] Column 4 is a weighted stabilized saving based on the differential between the alternative cost of producing comparable power by other means and our assumed cost of producing nuclear power (Table I, column 1). The weights are based upon Mayer's 1965 distribution of projected costs associated with the production of power in new steam generating plants. Mayer has estimated that large new steam plants will take, so to speak, 89.7 to 96.8 percent of the 'loss' to nuclear generation between 1956 and 1980. Savings for intermittent years between 1965, 1970, 1975, and 1980 are interpolated linearly.

[5] Assumes that in some magical way after 1967 the average cost of generating nuclear power in existing plants declines in such a way that the weighted saving expected to be associated with the construction of new plants (column 4) can be obtained for all nuclear power generated (column 1). This method of estimation overstates the true benefit in that once resources are committed to the construction of a particular type of reactor, it is impossible to take advantage of cost decreases which accrue to the fixed factors of production associated with new plants. Savings are biased further upwards since positive savings are also computed for plants constructed prior to 1968. For a discussion of the relation between stabilized costs as a function of the times when various plants come into initial operation and the cost of an individual plant in various time periods, see Davis and Roddis, *op. cit.*, p. 21-23.

(*) The saving for 1980 contains a capitalized return on 1,492.7 billion kwhr of nuclear power generated in periods beyond 1980. This method underestimates savings to the extent that either further cost decreases accrue as a result of funds invested in nuclear development prior to 1980 or additional savings are obtained by increasing nuclear capacity. These benefits, however, should be discounted very heavily for reasons of uncertainty and the fact that they will accrue only in the distant future; their magnitude would not be great and for practical purposes can be ignored. The rate of discount used to convert savings to 1967 dollars is again five percent.

[6] Assumes that the federal government will spend an average of 100 million dollars annually for research and development purposes exclusive of any subsidy that might be granted in support of uneconomic nuclear developments. Expenditures are assumed to be made over the period 1954-1980. As pointed out in the body of this paper expenditures through fiscal 1957 have already exceeded this rate. Recently a group of the nations' leading atomic reactor experts advised Congress that the Administration's program for atomic-power development is inadequate. The experts said that more government effort and money is required if nuclear-power development is to meet domestic and international needs. See, John W. Finney, "Reactor Experts Ask Rise in U. S. Atom Power Funds," *The New York Times*, January 12, 1958, p. 1, 63. In light of this climate of opinion, it does not seem likely that Congress will reduce appropriations for nuclear-power research and development within the foreseeable future.

should be pointed out that there is an analysis by Ullmann which entails even more favorable assumptions than the ones used in this paper; his assumptions, however, probably stack the deck in favor of nuclear development.)[10]

Some of the major assumptions involved in the benefit-cost analysis contained in this paper will be discussed in the remainder of this section. One of the problems in analyzing projected cost declines is that the projectors of declines usually fail to relate their projections to an assumed level of research and development expenditure. Ideally, cost declines should be directly related to expenditure levels; the economic problem would then be to choose the cost decline function and the implied expenditure level (optimally divided between research and development inputs so as to minimize expected costs) that would maximize net savings from nuclear development.

As is the case with most empirical work, one must be content with less than perfect data; it should be recognized, however, that a weakness of the approach used in this paper is that the method of estimating the costs of obtaining economic nuclear power is more independent of projected cost declines than is theoretically to be desired.[11] The reader should feel free to compare the cost decline assumed in Table II with the projections of others in Table I and to

check the reasonableness of assumed expenditure outlays for the development of economic nuclear power with other sources that might be available. The empirical assumptions in Table II are sufficiently explicit that anyone familiar with discount tables can make adjustments thought to be appropriate.

A few words should perhaps be said with regard to research expenditures. Only those of the federal government are included in the analysis contained in Table II. The 100-million-dollar annual expenditure assumed is conservative considering past expenditures,[12] the possible consequences of Sputnik, statements by leading reactor experts, and the politics which surround the public power question. Private research expenditures not covered by reactor sales are ignored. These may prove to be high owing to the competition which exists in the field and the fact that industry made sizeable investment in research related to civil power development on the basis of expectations which may prove to be unduly optimistic. The institutional framework surrounding the development of nuclear power is such that private industry, at best, can reasonably expect to capture only a small fraction of the returns from the development of economic nuclear power—these only in the short run and, if at all, only in the distant future when nuclear power becomes competitive with conventional power. From the point of view of reactor company stockholders the utopian dream of cheap nuclear power may have worked to their financial disadvantage.

With respect to estimating the cost of producing uneconomic power, a relatively short developmental period is

---

[10] For that matter, his assumptions do not even appear to be consistent when one stops to consider that nuclear power is bound to remain more capital intensive than conventional steam unless a way is unexpectedly discovered to circumvent the steam cycle. Real economies are generally predicated on the basis of fuel savings. Ullmann appears to predicate real economies on the capital side as well. "Economics of Nuclear Power," *Science* April 4, 1957, 739-43; letters of Kolin and Barnea, *ibid.*, July 4, 1958, 94-96.

[11] Research expenditures are estimated entirely independently of the assumed cost decline functions. With respect to the social and private cost of producing uneconomic nuclear power, one really does not know whether the methods used to forecast the growth of nuclear capacity during the developmental period result in a power output that is consistent with the attainment of the assumed cost decline.

[12] Teitelbaum and Mullenback have estimated that the development expenditures for civilian purposes are on the order of $150 million annually and rising rapidly. "The Development of Nuclear Energy," *Federal Expenditure Policy for Economic Growth and Stability, op. cit.*, p. 724.

assumed during which the larger and more efficient utilities experiment with nuclear power. The period 1954-1967 is suggested by the projection of nuclear power costs made by W. Kenneth Davis, Director, and Louis H. Roddis, Jr., Deputy Director, Division of Reactor Development, United States Atomic Energy Commission. These gentlemen envision an expansion phase in the growth of nuclear capacity after 1967 as nuclear power becomes obviously competitive with conventional sources. It should be pointed out that their projection is more optimistic than the tough minded analysis of Karl Mayer, which assumes that the growth of nuclear capacity will be only in response to economic forces.[13] If Mayer's projection of nuclear expansion were used instead of the Davis-Roddis projection, the resulting benefit-cost ratio would be lowered from (6.33 to 1.00) derived in Table II, to (2.34 to 1.00).[14]

It seems intuitively plausible that only the larger and more efficient utilities or groups of utilities will be able to command the resources necessary to experiment with nuclear power, spread the cost of uneconomic power to a large number of consumers, and be able (hopefully) to profit from the experience by adding additional capacity when economic nuclear power is finally achieved. Small high cost utilities have every reason to wait until there is assurance that eco-

nomic nuclear power has been achieved. The reasoning underlying the assumption of seven-mill conventional power during the developmental period could, however, be checked by looking at the generating costs of utilities applying for assistance under the Atomic Energy Commission's power reactor demonstration program.[15]

There does exist a difficulty in interpreting precisely what is implied by projected cost declines since the projections as a rule fail to distinguish between fixed and variable costs and the differential impact of technology on committed and uncommitted resources with respect to time. If we assume that the cost decline projections represent stabilized average costs of generating power over the life of the plant, with proper discounts given to both the high initial operating cost of starting up a reactor for the first time and any technological decline in variable costs that might occur with respect to time (this is how I interpret the Davis-Roddis projections), a short discount period, such as the 20-year amortization period assessed in Table II biases the resulting B/C ratio upwards. If one interprets the cost projections as making no allowance for changes in operating expenses over the life of the plant, then a short amortization period can be thought of as a compensation for pecuniary and technological economies obtained from a given reactor over time.[16]

In order to avoid the possibility of an unfavorable bias which could result from different interpretations as to what is implied by cost decline projections, the full value of the cost decline in each year

---

[13] Their estimate is larger than Mayer's estimate for the following reasons. Total electric energy consumption is forecasted to be higher, reflecting greater optimism with respect to the future use of residential electricity. Mayer has noted that the total difference between the two aggregate projections of electric energy consumed for 1970 is equal to about the difference in predicted residential use in 1970. Mayer, op. cit., p. 45. The Davis-Roddis estimates assume 70-80% load factors while the Mayer projections are based upon a 50% load factor for nuclear development. Finally the Davis-Roddis estimates assume larger increases in the alternative cost of producing power by conventional means. These differences in assumptions lead to an implicit forecast of nuclear-power output in 1980 nearly four times greater than the Mayer estimate.

[14] The calculations underlying this estimate can be obtained from the Office of Agricultural Economics, University of Chicago.

[15] While a check of this assumption is beyond the scope and resources of this study, it ought to be made. Clearly the net loss associated with the production of experimental power would be less, other things being equal, if it were institutionally possible to conduct the experiments in areas with relatively high conventional power costs.

[16] For further discussions of what is implied in cost decline projections see: Davis and Roddis, op. cit., p. 21-23; Tybout, op. cit., p. 758-80.

is assigned to both existing and new plants when estimating savings or benefits.[17] This has the effect of biasing upwards aggregate savings since existing plants cannot in fact take advantage of design economies which are surely allowed for in projecting cost declines. A further bias is obtained by imputing after 1967 savings on plants constructed during the developmental period.

Comparative Rates of Return

Ideally, one would want to compare the expected returns from developing nuclear power for civil purposes with the prospective returns from other technological innovations within the energy fields. While the benefits from harnessing solar energy or developing a commercially feasible method of liquifying or gasifying coal might be greater than the returns from nuclear power,[18] reliable data do not exist at the present time for making comparable statistical comparisons. Even within the atomic energy field, it is not at all clear that additional

---

[17] An interesting theoretical question can be raised as to whether the "savings principle" overestimates or underestimates the benefits from nuclear development. If demand is projected into the future without making allowances for increased consumption resulting from lower priced nuclear power, the method understates the true benefit; if allowances are made for increased consumption, the method overstates the true benefit since savings are computed for power that would not be consumed at prices prevailing in the absence of nuclear power. While this sort of bias can be of considerable importance in the case of navigation (Ed Renshaw, "The Measurement of the Benefits from Public Investment in Navigation Projects," *The American Economic Review*, September 1957, p. 652-62), it is of only minor importance in the case of nuclear development. Nuclear and conventional power can be considered almost perfect substitutes. The increased consumption resulting from a lower price would not be great owing to an inelastic demand for power and the fact that generating cost is often a small fraction of the total price of power; it is even less consequential when one stops to consider that the projected nuclear demand in Tables 2 and 3 differs by a factor of nearly four, as a result of other assumptions.

[18] E. Finley Carter, Director, Stanford Research Institute has suggested that the potentiality of western coal is not being realized because of unsolved technical and economic problems centering around getting the coal to market or converting it at the mine or elsewhere into electrical, gas, or liquid energy. In commenting further he maintains this is one field of research which deserved a "high priority." *Federal Expenditure Policy for Economic Growth and Stability, op. cit.*, p. 1154.

investment in civil power will yield the highest returns. Addresses by AEC Commissioner Libby have reported that hundreds of millions of dollars have already been saved by industrial applications of isotopes. In the long run, returns might be greater if more resources were directed towards basic, general, and less directed lines of fundamental inquiry.

Agriculture is one industry in which a great deal of careful work has been done in an attempt to isolate the savings from new techniques of production. Drawing on the work of Schultz, Jenkins, Mighell, and others, Griliches has made estimates of the social returns from the development and adaptation of hybrid corn, the prospective returns from a hybrid sorghum, and the returns from total public and private expenditures on agricultural research.[19] His ratio of benefits to costs for hybrid corn is 155 to 1; for hybrid sorghum, 50 to 1; for agriculture as a whole he obtains an upper limit of 34 to 1 and a lower limit of 7 to 1. These estimates can be compared with the benefit-cost ratio of 6.33 to 1.00 computed from Table II. It should be noted, however, that Griliches adopted assumptions which would bias downwards the resulting benefit-cost ratio, while the assumptions adopted in this paper tend to bias the ratio upwards.

The returns he obtains for agriculture are comparable to estimates made by R. H. Ewell for the economy as a whole (100 to 200 percent per year per dollar spent on "research and development" which would imply benefit-cost ratios ranging from 20 to 1, to 40 to 1), and figures quoted by major industrial companies on their returns on research.[20]

---

[19] Griliches, *op. cit.*, A simple translation between the rate of return as Griliches computes it and the benefit-cost ratio used in this paper can be made by multiplying the benefit-cost ratio by a factor of five—the rate of interest used in discounting—or by multiplying the rate of return by the reciprocal of five.

[20] R. H. Ewell, "Role of Research in Economic Growth," *Chemical and Engineering News*, 1955, 298-304.

The comparisons would seem to indicate that the returns from developing fission power are not expected to be as high in the United States as the returns from research in other areas. This is not to contend that further research and development in the area of achieving economic fission power would constitute a bad investment but rather to point out that the justification for an accelerated rate of public and private expenditures aimed at quickly attaining that end must be justified on other grounds. (The civil power program is at the present time absorbing resources equivalent to from one-third to one-half of the total federal expenditure for non-security research.)

Over the long haul, economic fission power could be obtained primarily as a by-product of more urgent research in military reactor technology or at a more leisurely and less costly rate of development.[21] Reports of the Paley Commission and others are reassuring in that the United States need fear no shortage of fossil fuels to meet an expanding energy output within the foreseeable future.[22] In the long run the development of the fusion principle could conceivably make the fission principle obsolete.[23]

[21] "A high priority for atomic power might justify the present extent of demonstration program aid for the relatively small participating group. On the other hand, a gradual development of atomic power might look toward the training of a larger number of participating firms at less cost for each, and with more time to cultivate competitive relationships among power plant designers and suppliers. Similarly, in AEC's own civil power research, a more leisurely approach, exploring one technological finding after another, might lead to the same end results at lower total cost than the present multifront attack." Tybout, op. cit., p. 795.

[22] Tybout, op. cit., pp. 747-53.

[23] Fission power plants are similar to conventional power plants except that steam is generated in the nuclear reactor, or in a heat exchanged by a fluid which has in turn received its heat from the reactor; if relative economies are to be obtained, they must come from a reduction in the cost of fuel rather than a smaller capital outlay. The fusion principle offers hope of converting energy directly into electricity, thus holding the potentiality of economies on the capital side. A further advantage of the fusion principle is an elimination or lessening of the radiation hazard. Advantages are sometimes cited with respect to problems associated with the control of the reaction.

If a priority is to be given at this time to the development of nuclear power, it must be done on grounds such as: complementarity between civil and military reactor technology; indirect benefits or by-products that might result from directing research effort towards the attainment of economic nuclear power; savings that will accrue to our allies and to other countries with relatively high power costs; or national prestige. While any of these reasons might be sufficient to justify accelerating the civil power program, they should, nevertheless, be subjected to the careful scrutiny of expert opinion and compared with alternative ways of obtaining the same ends. Otherwise the public will have little assurance that the objectives of the Atomic Energy Act of 1954 are being lived up to. Broadly, these are: (a) The development, use, and control of atomic energy shall be directed so as to make the maximum contribution to the general welfare, subject at all times to the paramount objective of making the maximum contribution to the common defense and security. (b) The development, use, and control of atomic energy shall be directed so as to promote world peace, improve the general welfare, increase the standard of living, and strengthen free competition in private enterprise.

Given that there exists a willingness and a desire to promote the development of economic nuclear power as a part of overall national and international policy, the economic problem is two-fold. First, there is the question of efficiency. If, in order to obtain our objective, we must first construct experimental reactors which will produce power that will not be competitive with conventional sources, what kind of institutional arrangements can be made to minimize the cost of uneconomic power? Secondly, do the exist-

ing institutional arrangements distribute the cost of nuclear development equitably among the beneficiaries and the public at large?

While the detailed cost analysis that would be necessary to answer these questions is beyond the scope of this paper, doubt has been raised as to whether the existing Atomic Energy Commission program strives to minimize the costs of producing uneconomic power. Other things being equal, it would be highly desirable to experiment with nuclear power in those areas of the United States and the free world where conventional power is most costly.[24]   Further doubt can be raised as to the fairness of shifting a substantial part of the cost of uneconomic nuclear power on to particular power consumers rather than resort to direct subsidies borne by the public at large.

----

[24] Mayer has suggested that consideration be given to the granting of subsidies to United States firms engaged in nuclear energy activities in high-fuel-cost areas overseas. *Federal Expenditure Policy for Economic Growth and Stability*, *op. cit.*, p. 714. For its part in atoms-for-peace work, Atomic Energy Commission requested a budget authorization of approximately $7.7 million. Tybout, *op. cit.*, p. 769. This relatively small amount can and should be compared to the billions in direct economic aid given to our allies and the free world, and to billions which will undoubtedly be spent on a domestic civil power program.

Los Angeles: Urban Prototype†

By ARTHUR L. GREY, JR. *

Bellweather of Change

IN the nation at large Los Angeles has long been considered unique. The picture is in part a myth, nurtured by local boosters for self-interested and romantic reasons and by people in other parts of the country who may take consolation in this delusion. A national magazine has devoted an issue to what it called, "The Astounding World of Los Angeles." This popular characterization obscures the larger significance that Los Angeles has for all cities. The fact is that the formerly distinctive features of this area foretold what have become widespread urban trends.

One reason for this is that Los Angeles' rapid and continuing growth has provided a fertile and demonstrative field for innovations of all sorts. Bogue has pointed out that "the Los Angeles area has grown faster during each decade since 1900 than the Standard Metropolitan Areas of its size and class in the Nation."[1] From 1900 to 1950 Los Angeles' population increased 2,199 percent while during the same period the average growth for the fifty-seven principal Standard Metropolitan Areas in 1900 (including Los Angeles) was 160 percent.[2] Under these circumstances physical changes which might be of minor proportions in other localities have readily emerged as the prevailing mode in Los Angeles.

The Means of Passenger Transportation

Another consideration is the fact that Los Angeles is pre-eminently a product of the automobile age. If we were to identify any single variable as being of the most significance in determining the essential points of difference among our larger cities it would be the dominant means of transportation during the period of major population growth. In California this point is readily demonstrated by the contrast between San Francisco and Los Angeles. By 1880 San Francisco was the ninth largest city in the country with a population of 229,000 or over one-one-fourth its present population. Los Angeles did not reach this size until after 1900 and its population then constituted only about four percent of what it is today.[3] In San Francisco the major elements of existing development were conditioned by rail transportation, which meant that growth was vertical. The average population density in 1950 was approximately 16,000 persons per square mile. That the 1950 population density in the city of Los Angeles was only about one-fourth as great as that in San Francisco, or about 4,000 persons per square mile, must be accounted for to a significant extent by the contrast in methods of transportation.

The salient features that distinguish the physical development of Los Angeles had their genesis to a large extent in the decade of the 1920's. This was a period of rapid population and economic growth for the area. Professor Robert A. Gordon

[1] Donald J. Bogue, *Population Growth in Standard Metropolitan Areas, 1900-1950; with An Explanatory Analysis of Urbanized Areas* (Washington, D. C.: Housing and Home Finance Agency, December 1953), p. 56.

[2] *Ibid.*

[3] A few years prior to this date a San Francisco observer commented: "It is hardly too much to say that the modern horsecar is among the most indispensable conditions of modern metropolitan growth . . . . [The] horsecar virtually fixes the ultimate limits of suburban growth . . . ." *San Francisco City Directory for the Year Beginning March 1875* (San Francisco, California: Henry G. Langley), p. 20.

in his study of the business cycle during the interwar boom has concluded that much of the prosperity of the 1920's ultimately can be traced to the mass production of the automobile and its other effects in street and highway building, petroleum, tires and rubber, and in suburban residential growth.[4] In the 1920's some 418,000 new homes were built in Los Angeles city. This more than doubled the existing housing stock. Most significantly, by 1930 the percentage of the country's automobiles within Los Angeles County was nearly twice the proportion which the area contained of the country's total population.[5]

The question arises of why the Los Angeles area had an unusually high ratio of automobile ownership; of why automobile usage developed on a more generous scale earlier than elsewhere. Among the answers are such factors as the minimum of restraint imposed upon the direction or extent of new growth by topographical conditions and the mild climate which made all-weather operation of the automobile feasible. Climate also made a lower standard of public services practicable and encouraged more dispersed development dependent upon the automobile than occurred elsewhere. The early use of the automobile was to a much greater extent recreational than is true today. Los Angeles was much vaunted for its scenic beauty before automobile development and population growth had consumed the open spaces.[6]

Finally, there was the antecedent development of streetcar lines which themselves fostered dispersed development in Los Angeles.[7] Street railway development in many localities furthered land subdivision and development, and in Los Angeles this was markedly the case. Henry Huntington was recognized to be the "greatest single land owner in Southern California" and in Southern California he developed the most extensive streetcar and interurban system in the world. Huntington's rail lines linked together and assured the success of many farflung townsites. The passenger facilities of this rail network are now almost completely dismembered and it is difficult today to imagine that mass transportation ever made an appreciable contribution to the making of modern Los Angeles. However:

"The effect of the railways in the development of southern California is fairly easy to trace. A sharp increase in population and assessed property valuation is evident in communities after their railways were built . . . . Considering the lack of exceptions and the fact that most towns without a railway were stunted, it seems fairly clear that electric transportation influenced the area's growth to a marked extent . . . ."[8]

"Few districts in the United States owe as much of their character to local rail transportation systems as does Southern California . . . . Few districts were similarly blessed with climate and topography which enabled them to take advantage of interurban development."[9]

[4] Robert A. Gordon, "Cyclical Experience in the Interwar Period: The Investment Boom of the 'Twenties," *Conference on Business Cycles* (New York, New York: National Bureau of Economic Research, Inc., 1951), p. 190.

[5] Los Angeles County in 1930 accounted for 1.8 percent of the nation's population, 3.5 percent of the registered automobiles.

[6] The following ecstatic comment appeared in a 1915 travel book: "Southern California is easily the motorist's paradise over all other places on this mundane sphere. It has more cars to the population—twice over—and they are in use a greater portion of the year than in any other section of similar size in the world and probably more outside cars are to be seen on its highways than in any other locality in

the United States. The matchless climate and ever-increasing mileage of fine roads, with the endless array of places worth visiting, insure the maximum of service and pleasure to the fortunate owner of a car, regardless of its name plate . . . ." Thomas D. Murphy, *On Sunset Highways* (Boston, Massachusetts: Page, 1915), pp. 3-4.

[7] Glen Dumke, "Early Interurban Transportation in the Los Angeles Area," *Southern California Historical Quarterly*, September 1940, p. 132. Nevertheless, the mistake has been made concerning Los Angeles of assuming that "since no street railway system existed, the pressure for more roads for auto traffic became very strong, tending to obscure the need for public transit." Fred K. Vigman, *The Crisis of Our Cities* (Washington, D. C.: Public Affairs Press, 1955), p. 140.

[8] *Ibid.*, p. 144.

[9] *Ibid.*, p. 131.

The point is, then, that Los Angeles started out like other communities (at least in respect to those considerations that are important to our present analysis) but that at an earlier date than in any other large metropolitan area, and in very many of the smaller ones, the automobile became a dominant element in local passenger movements. In this respect as much as in others, the trite "colossal" of the oldtime Hollywood movie moguls characterizes the scope of the developments that have taken place in Los Angeles.

Decentralization: Residential and Commercial

The swing in new construction to the single-family house which dates from the early 1930's doubtless can be traced to a coterie of forces, including the influence of the federal programs in the home financing field and rising levels of personal incomes. A decisive factor also has been the trend to the suburbs made possible by the automobile, which Gordon identifies as already an important factor in suburbanization in the '20's.[10] In Los Angeles, vast new suburbs have grown up without practically any public transportation whatsoever.

Suburbanization developed to a unique degree in Los Angeles. Bogue points this out when he said: "The Los Angeles Standard Metropolitan Area is much more decentralized (in the sense that a larger proportion of its total population resides outside the central city in the suburban ring) than any other SMA's of its size and region."

The tract house was the symbol of the move to the suburbs which resumed on a greater scale than ever before in the latter 1940's after the interruptions of depression and war. A more recent phenomenon has been the development of the suburban shopping center as the

logical companion of the outward migration of people. In Los Angeles this trend has been carried to near saturation for present levels of population and purchasing power. Besides the seemingly numberless smaller centers the main attraction of which is a food store or perhaps a "junior" department store, there are more than a score of so-called "regional" shopping centers which account for 38 department stores opened between 1945 and 1958.

A recent survey shows that from 1948 to 1954 retail sales (in current dollars) fell by 6.7 percent in downtown Los Angeles as compared with a drop of under two percent in the Central Business Districts of all cities of 1,000,000 or more population. Some other large cities are experiencing impressive sales drops—in downtown Philadelphia the decline was 7.2 percent and in Detroit 11.2 percent in the 1948-54 period, while in the Metropolitan Areas in the 500,000-to-1,000,000-population class, sales declines of 10.2 percent and 16.0 percent took place in St. Louis and Pittsburgh respectively.[11]

Actually, the full impact of shopping decentralization upon downtown Los Angeles has been cushioned by the very rapid rate of local population growth. The magnitude of the shift away from downtown is disclosed more fully in the fact that during the same period the increase in retail sales throughout the whole Los Angeles Metropolitan Area was over 50 percent, which is considerably more than in any other in the country and may be compared with the 30.6 percent increase in Philadelphia and the 39.4 percent increase in Detroit.

Retail decentralization has had telling effects over a much longer time in Los Angeles than elsewhere. For nearly 25

---

[10] Gordon, "Cyclical Experience in the Interwar Period," *loc. cit.*

[11] Murray D. Dessel, "Central Business Districts and Their Metropolitan Areas," United States Department of Commerce, *Area Trend Series*, November 1957.

years sales totals for downtown department stores have been following a downward trend and in the period 1946 to1957 the decline has amounted to 30 percent.[12] This drop is much more impressive when one considers that these figures take no account of purchasing power changes although the post World-War-II period has been one of sharp inflationary trends. Thus Los Angeles is in the vanguard of the decentralizing trend.

The author is mindful of a contrary opinion expressed by Francis Bello that "every city" is not necessarily destined "to look like Los Angeles. Los Angeles had no vital central core to begin with." Bello exaggerates the fairly subtle differences which distinguished Los Angeles at an earlier date when he asserts that "the core *never* developed, despite a fine electric railway system in the early years, because the automobile arrived too soon and splattered the city over the countryside."[13] [Italics added] In point of fact, downtown Los Angeles has been an important center and only since World War II has the decrease in its importance been precipitous, as figures on department store sales, for example, indicate. In the space of the 12 years beginning with 1946, the downtown share of the Metropolitan Area sales volume has been more than cut in half, from 44.8 percent to 19.6 percent.[14] The former figure may be compared with 1954 figures of 39.8 percent for Detroit, 52.4 percent for Milwaukee, and 47.8 percent for Buffalo, as disclosed by the United States Census of Business.[15] Going back to 1930 when 60.1 percent of the population of the County lived within Los Angeles city, 87.9 percent of the department store sales in the County were transacted there.[16] While figures are not available separately for the downtown in that year, the location of the stores as shown in the city directory and knowledge of their approximate sizes indicate that not less than two-thirds of all department store sales within Los Angeles County then were attributable to establishments in the Los Angeles central business district.

Political Fractionalization

Students of government and others decry the lack of administrative machinery in Los Angeles to handle effectively many problems which are area-wide in scope. In a five-county area of which Los Angeles is the hub, there are about 1,000 governmental units including over 110 cities. In Los Angeles County there are 63 cities; and even so there are over a million people living in unincorporated areas of the County. During the past three and one-half years the formation of new cities has occurred at an epidemic rate; 15 incorporations have actually been completed and in other areas steps toward incorporation have been made. As compared with the 17 which have incorporated, only six incorporation proposals which have been voted upon have failed to carry. This is sharply in contrast with the record in the years 1930-1954 when all but two incorporation attempts failed.

The concomitants of this division of responsibility are very well recognized: inadequate and badly distributed tax bases; substandard development in newer and less competitive areas; and dilution and loss of political leadership as members of the middle class move to the suburbs and lose their citizenship in the

[12] Security-First National Bank of Los Angeles, *Monthly Summary of Business Conditions*, December 1957.
[13] "The City and the Car," *The Exploding Metropolis* (Garden City, New York: Doubleday & Company, Inc., 1958), p. 57, reprinted from *Fortune*, October 1957.
[14] Research Department, Security-First National Bank of Los Angeles.
[15] United States Bureau of the Census, *U. S. Census of Business: 1954*, Central Business District Statistics Bulletins CBD-45, -2, and -37, Washington, D.C., 1956.
[16] United States Bureau of the Census, *Fifteenth Census of the United States, 1930*, Distribution, Vol. 1, pt. 1, p. 260.

older cities. As to the causes of the most recent rash of incorporations, one resume comments:

"The greatest single factor in the incorporation outbreak is the desire of certain communities to preserve and others to change the land use pattern. The rush for land has caused some of the well established unincorporated areas to fear changes in the county zoning ordinances or proposals to annex adjacent cities. Others complain that their growth is stifled by the present land use pattern allegedly held static by county ordinances."[17]

Paradoxically, Los Angeles was a pioneer in city planning in this country. The city of Los Angeles has had a planning department with a large staff for many years and in name at least the County has recognized the scope of the need for a long time—the Los Angeles Regional Planning Commission was created in 1923.

On the other hand, the solution of the Los Angeles governmental problem is not to be sought simply in a reversal of this trend and the creation of larger political units. The city of Los Angeles is in many ways too big and too centralized. The San Fernando Valley is a vast area of about 212 square miles within the city limits of Los Angeles. Most of this area was agricultural land when World War II ended. Twelve or fifteen years ago Gordon Jenkins made the Valley famous in a popular song as an idyllic sort of place, but more recently it has become the prototype of suburbanization and, to its critics, a sort of nightmare with miles of tract developments, many without such elemental conveniences of city living as street lighting, sidewalks, storm drains, sanitary sewers, or parks. "The planing of land uses in the San Fernando Valley is an extreme but not isolated example of the problems facing planners in controll-

ing undeveloped but rapidly growing areas."[18] Certainly the Valley is little different in character from other territories in the southern part of the County or in the eastern San Gabriel Valley which are governed by numerous separate municipalities.

The smaller cities are wide open to criticism from the standpoint of planning and governmental efficiency but large size, too, can be a disadvantage. The most remote part of the San Fernando Valley is 35 miles from the Los Angeles City Hall but many residents have felt, no doubt not always with complete justification, it might as well be 350 miles. The solution to be found in the governmental fields is a political structure which will match the scope of the problem but allow effective participation and leave enough autonomy to deal with purely local (in contrast to metropolitan area or regional) matters.

Social Implications

The Gordian knot of race relations outside the South at the present time may be the matter of housing.[19] Even equality of opportunity for employment may prove much more amenable to positive action in the short run than will the integration of residential neighborhoods. Dr. Morton Grodzins recently told an annual conference of the National Association of Housing and Redevelopment Officials that "racial schism is the biggest impediment to improved housing" in our

[17] *BGR Observer*, Bureau of Governmental Research, University of California at Los Angeles, June 1957.

[18] Fred E. Case and James M. Gillies, "Some Aspects of Land Planning: The San Fernando Valley Case," *The Appraisal Journal*, May 1954, p. 28. These writers term the San Fernando Valley an example of "flexible planning" (*ibid.*) because of the numerous instances in which the Planning Commission set aside previously prepared plans when confronted by the pressures of conflicting private speculative interests.

[19] The Commission on Race and Housing, *Where Shall We Live? Conclusions from a Three-Year Study of Discrimination in Housing* (Berkeley, California: University of California Press, 1958), p. 3 states that: "Housing and residence . . . . have proved probably the most resistant of all fields to demands for equal treatment."

big cities generally and that the big cities are being turned into "lower class, largely Negro slums" while they are being surrounded by a suburban "ring that is composed generally of middle class, white, Protestant and Republican" home-dwellers.[20]

Los Angeles is recognized as having made many notable strides in race relations, but in the matter of housing it has failed to evolve an integrated pattern. The old residential portions of Los Angeles and the other long established cities of the area, as housing has become obsolete or rundown, have been tenanted by the Negro and Mexican minorities but the new suburbs have been virtually impenetrable to these groups.

The ironic role of public action through the policies instituted by the federal housing agencies in rationalizing and perfecting a code for discrimination in housing with respect to race and ethnic origin is well-known.[21] The building of homes uniform as to square feet of interior space and selling prices tends to concentrate populations homogeneous in other respects—family size, income, age composition, and occupation. The rules prescribed in this regard by the public agencies acting in the cause of minimizing credit risks were not of such great importance initially as in their application later to the great postwar housing subdivisions.[22]

The legally enforced use of the protective covenant, to which the Federal Housing Administration gave much credence, has of course been struck down by the courts as it pertains to the restriction of occupancy on the basis of race, color, and national origin. However, this has not disenchanted the housing agency with its use in other respects and, although the Manual no longer makes specific reference to these considerations in the criteria of appraisal valuations, as recently as 1955 the doctrine of homogeneity was still in force, including a criterion of "Compatibility Among the Neighborhood Occupants."[23] This must have been tacitly understood by many schooled in the old ways to include the absence of minority occupancy which, after all, is commonly considered to be an important standard of neighborhood harmony.

To some observers the sameness with respect to race and the other characteristics is the worst feature of the postwar mass housing developments. Such communities may militate in favor of conformity and perhaps create ignorance and intolerance by insulating the standardized dweller in a standarized house

[20] Associated Press dispatch, October 14, 1958.

[21] One of the first commentaries to draw attention to these facts was the study by Myrdal and related investigations of the position of the Negro in American society undertaken through the auspices of the Carnegie Foundation. See Gunnar Myrdal, *An American Dilemma* (New York, New York: Harper & Brothers, 1944), pp. 348-350, 625; and Richard Sterner and Associates, *The Negro's Share* (Harper & Brothers, 1943), pp. 312-316.

[22] The following is indicative of the sociology, outside the more controverted field of race, etc., which the Federal Housing Administration at one time dispensed to its appraisers: "At the bottom of the employment scale, a large percentage of the employed population of a city is found working in the capacity of laborers. A member of this group has little or no individuality, because any one person can usually perform the work as well as any other . . . The next group encountered . . . . includes the semi-skilled

mechanics and lower-grade clerks. The members of this group exhibit some individuality, because they have demonstrated through knowledge or ability a degree of superiority over laborers. Further up the scale are encountered mechanics and office workers. The capacity of members of this group is apparent, because they have shown the ability to rise to a point above the large groups below them. Next in the scale are found the foremen, chief clerks, and others of like standing in the employment world. These are usually well established people. Above this group are the junior executives. It is considered that this group presents, for the purpose of determining mortgage risk, the best income characteristics in the entire employment field. The usefulness, versatility, and ability of this group have been firmly established. Should they lose their present employment, it is usually possible for them to find ready employment in a similar line of activity . . . . Should it become necessary for members of this group to seek employment in lower strata of the employment scale, they can readily replace those below them by reason of their demonstrated superior ability and usefullness." Federal Housing Administration, *Underwriting Manual, 1936* (Washington, D. C.: United States Government Printing Office), Sec. 219 [2].

[23] *Ibid., 1955,* Sec. 1315.

from contact with social groups different from his own. Because there may be something highly inimical to further progress in race relations and in human relations generally in such set living patterns, the process and the standards which enforce this sameness need evaluation, as do the consequences. On a greater scale than anywhere else the uniform neighborhood, as the direct result of the rating criteria followed by the Federal Housing Administration, and also by the Veterans Administration, was propagated in Southern California. Nearly nine percent of the dwelling units erected in the entire nation in the years 1946-49 inclusive were built in the Los Angeles Standard Metropolitan Area and for the period 1950-58 the proportion was about 7.8 percent.[24]

Interestingly enough, California legislated virtually out of existence the use of public housing in redevelopment areas which might give them in small degree the saving grace of heterogeneous population by passing a state constitutional amendment requiring local referendum approval of such housing. In Los Angeles, where sentiment for the amendment originated, this has assured quite completely that no public housing will be built.

Disorganized Growth and Destruction of Open Space

Former United States Housing Administrator Cole comments that "for the last ten years or more, people have been flooding out of the cities into the countryside—a modern gold rush to find a spot of sunshine. But where is the country? Unless you go out 50 miles or more, the landscape now is scattered with septic tanks, filling stations and super-highway strips."[25] Lewis Mumford complains that in the environs of New York City the countryside has been despoiled by the progressive advance on a continuous front of housing development and "random" highway building which has had as its "chief result" the extension of the congestion of midtown New York to outlying areas and which "has added to the disorganization and frustration of urban life . . . ."[26]

The urban problems confronting Los Angeles are those facing the nation. William Whyte says that an air trip from neighboring San Bernardino to Los Angeles can provide an "unnerving lesson in man's infinite capacity to mess up his environment."[27]

In the quest for living space we have created a pattern which has unnecessarily removed open space further and further from those who dwell in the cities. Further "channels of escape," Mumford states, must be provided "at greater distances from that once so admirable center; finally, intensified congestion both in the original center and in the suburb . . . . wipes out the social assets of the city and the rural assets of the country." He terms this process "environmental impoverishment."[28] William Zeckendorf says that in the Angelenos' "urge to escape they have created a condition from which there is no escape . . . . There is no out of town in Los Angeles."[29]

For many years Los Angeles County had such a commanding lead in agricultural production that the statistics

[24] Los Angeles data from Research Department, Security-First National Bank of Los Angeles; national data from Housing and Home Finance Agency, *Housing in the United States, A Graphic Presentation* (Washington, D. C.: United States Government Printing Office, 1956),, p. 66; and Bureau of Labor Statistics, *Monthly Labor Review.*

[25] Address by Albert M. Cole, United States Housing Administrator, to the Municipal Forum of New York, April 11, 1958.

[26] Lewis Mumford, *From the Ground Up* (New York, New York: Harvest Books, Harcourt, Brace and Company, no date) p. 233.

[27] William H. Whyte, Jr., "Urban Sprawl," *The Exploding Metropolis*, p. 115, reprinted from *Fortune*, January 1958.

[28] Mumford, *op. cit.*, pp. 237-38.

[29] William Zeckendorf, "Fluid Suburbia," *Yale Review*, September 1958, p. 28.

failed to reflect fully the rate at which citrus groves, truck gardens, and dairy pasture were being converted into suburban lots. Recently, however, the county has been nudged from its position as first-ranking county in value of agricultural production in the nation and has fallen to third place. To some protagonists of urban "progress" this process of substitution may appear clearly as a gain, but this is to place no value on the quality of urban life, only on its quantity. The incursions upon open space have not been limited to fertile cropland; people and the banal artifacts of the "exurban" life have taken root in disarray in the far corners of Southern California. *Sunset*, ("The Magazine of Western Living") for April 1958, describes as follows a trip to the Joshua Tree National Monument, which—at its nearest boundary—is about 100 air miles east of the center of Los Angeles:

"For the next 43 miles [from Whitewater in the San Gorgonio Pass] to Twenty-nine Palms, you get a powerful demonstration as to how time is running out for California desert country that isn't preserved as dedicated park land: Where yesterday the fabulous Devil's Garden and the purple vale of Morongo lured painters and poets back again and again, the views aren't pretty any more. Spread across the landscape is a gridiron of "desert homesites" strewn with minimum size cabins as far as you can see, and there's a parade of billboards promising more such Valhallas [sic] around every bend in the road. Respite finally comes when you enter the Monument."

Industrial Dispersion

A concomitant of residential and commercial decentralization is, of course, a similar development in manufacturing. No attempt will be made here to measure the extent of this in Los Angeles in comparison with other communities, although some qualitative judgments will be ventured.

The various factors cited as determining the choice of location by manufacturing firms suggest a number of reasons for recent decentralization, *viz.*, availability of sites of suitable size, lower land prices, lower local taxes, zoning and other regulations (in Los Angeles this would include smog controls), more tractable labor. These differential opportunities to business are more readily exploited, depending upon the extent to which the level of development of the highway and street system minimizes space friction. Whether or not such dispersion is economically desirable on broader social grounds as well as to the particular units making their locational decisions, depends upon the totals of costs and benefits to business and sustained indirectly by the society. The development of greater mobility of the factors of production through transportation improvements and the consequent widening of choice of location by various enterprises may or may not prove a social advantage. An analysis of industrial locational decisions within the Los Angeles area indicates that in most cases the disadvantages of the selected site were ignored at the time of decision or discovered only later. For all firms, around 20 percent reported subsequent dissatisfaction with their choice of site, the proportion being higher for smaller firms.[30]

The provision of transportation facilities affects and is affected by the location of industry. The future impact of the latter aspect has been considered by Ray Vernon:[31]

"As manufacturing pushes its way into the suburbs, there is a considerable likelihood that new patterns of freight movement will

[30] Hadley E. Smith, *A Review of Plant Location Theory and a Questionnaire Study of Plant Location in Los Angeles County* (unpublished M. A. thesis, Los Angeles, California: University of Southern California, 1954), p. 61.

[31] "Production and Distribution in the Large Metropolis," *Annals of the American Academy of Political and Social Science,* November 1957, pp. 28-29.

develop . . . . This would mean that urban areas would have to accommodate themselves to diffused cross-hauling from one suburb to the next, rather than to radial movements to and from the central city."

Vernon then points out that since the existing rail and highway freight systems of the larger metropolitan areas are predominantly radial in layout their inadequacy for dealing with diffused movement is "striking."

Los Angeles' freeway system will soon include connections which will facilitate circumferential intra-regional movement. As Vernon points out, if his hypothesis is correct the urban complex is destined to undergo further pronounced changes in structure and land use. Because of (1) the great extent to which industry in the Los Angeles area has already located in the suburbs and (2) the extent to which motor vehicular freight traffic is already possible because of freeway development, Vernon's forecast probably is nearest to fulfillment in Los Angeles. Indeed, although no comprehensive origin-destination survey exists for the Los Angeles area, there can be no doubt but that at present a very large part of the area's freight movements take place without involving the central core of the city at all.

General Observations and Conclusions

The reason for fixing attention upon Los Angeles is that developments which are yet only incipient in many other parts of the country have become the established pattern in Los Angeles so that to examine this pattern is to extrapolate a trend—not to be followed in every detail but demarcing the likely course of events in other metropolitan areas.

For a generation the face of urban America has been in the process of being remade on a new pattern and this development has been especially strong within the past prosperous decade. The

earliest significant signs of this trend appeared in Los Angeles—in such inconsequential things as the drive-in theater, the drive-in restaurant, the drive-in bank, the single-family "ranch" house, the attenuation of commercial areas along all principal streets, the atrophy of the public transportation system, the supermarket, the shopping center, branch store retailing, motels, freeways, the large scale residential tracts, municipal annexations and satelite communities—to mention a few of the features which spell Los Angeles to the rest of the country and which have since become elements in the development of other localities.

Recently there has been great interest in the phenomenon whereby the great majority of the population of the United States is becoming concentrated in 15 or so vast belts of continuous urban development of moderate population density.[32] One of these so-called "strip cities" is the area from eastern Massachusetts to northern Virginia along the Atlantic seaboard which is presently the subject of investigation by the Twentieth Century Fund.[33] Another is Los Angeles joined with Riverside and Long Beach, and reaching to San Diego. By 1970 the territory from Santa Barbara on the north to Baja, California is expected to coalesce into a single metropolitan area. A decade ago, Carey McWilliams pointed out that:[34]

"One of the smallest geographic regions in America, Southern California is today one of the most intensively developed . . . . No other region in America, write the geographers

[32] For example, see John M. Willem, "Interurbia is Here to Stay," *Business Horizons*, Spring 1958, pp. 25-32; or Christopher Tunnard, "America's Super Cities," *Harpers Magazine*, August 1958, pp. 59-65.

[33] For an identification of the area of the "Megalopolis" study, see *Annual Report of the Twentieth Century Fund*, 1957.

[34] *Southern California Country* (New York, New York: Duell, Sloan & Pierce, 1946), pp. 12-13.

White and Foscue,[35] is dominated by one city to the extent that Los Angeles, with its 450 square miles of territory, dominates Southern California . . . .

"For all practical purposes, it is a non-rural region; there are few strictly rural districts. In effect, Southern California constitutes a single metropolitan district which should be characterized as rurban: neither city nor country, but everywhere a mixture of both. Just as Southern California is the least rural of all regions in America, so, paradoxically, Los Angeles is the least citified of all the cities of America."

Harry Ashmore, in writing on the school integration problem in Little Rock, Arkansas, declared that something of New York has pervaded all American communities and that its influence is irresistably transforming our culture into a single, more homogeneous whole.[36] In many respects New York is the model of American urbanism, but in certain respects it is an archetype, the epitome of urbanization rather than the norm toward which our cities are tending. The communities we are now in the process of evolving, the "rurban" strip cities, are distinctly different from the traditional city, which New York still symbolizes and which all other American cities reflected in varying degrees of imperfection a generation or two ago and are strongly influenced by the pattern developed during the past 40 years in the Los Angeles area which has increased its population six-fold.

Why is Los Angeles the prototype of decentralization and why have these features assumed importance first in this community? To be sure there are other, and probably unassociated, causes, but the most distinctive and determining consideration has without question been the unusual importance of the auto-

mobile. The discussion on the foregoing pages indicates that these ingredients disposed Los Angeles toward a high and early incidence of automobile usage: (1) The modest dispersion of growth and development which antedated widespread use of the automobile for the reasons cited. (2) The consequent receptiveness of the community to the introduction of automobile use. (3) The subsequent great population growth which has been accommodated by dispersed development primarily oriented to automobile transportation. (4) Events also show that the trend toward decentralization has been accelerated by the growing differential between the increments of new capital flowing, respectively, into the improvement of streets and the transit industry since the early 1900's.

This last fact has had much to do with the waning influence of mass transit facilities upon the course of city development and the increasing distress of central areas. Governor Furcolo told the opening session of the Massachusetts legislature in 1958 that one of that state's two most serious problems is mass transportation (the other is education). Yet no community or area, despite protestations over the transit situation, has done anything to redress the balance in transit and highway outlays. In New York where transit is now considered to have reached the crisis stage, the Interurban Rapid Transit has commenced a $100 million modernization program; at the same time, however, a two billion dollar highway-freeway program is progressing in the city's immediate environs. In Los Angeles the situation is simply farther advanced: road building got an earlier start and transit died an earlier death.

Originally, and before its high frequency of ownership precipitated traffic problems, the automobile afforded its

[35] Charles L. White and Edwin T. Foscue, *Regional Geography of Anglo-America* (New York, New York: Prentice-Hall, Inc., 1943), p. 735.
[36] Harry S. Ashmore, "A Southern Challenge and Epitaph for Dixie," *Life*, November 4, 1957, p. 146.

greatest advantages. So long as the parameters of the economy and social organization generally remained essentially unchanged, the automobile conferred new freedom on those who had cars. However, the advantages turned into compelling necessity. Instrumental in the growth in automobile ownership and usage after cars had become quite commonplace was the circumstance that, when decentralization reaches a certain stage, it engenders forces "that themselves accelerate its spread."[37] It becomes impossible to get along without the automobile in the new districts whose development has been predicated primarily upon automobile ownership and automobile inroads upon transit usage in the established areas lead to curtailments of service which force others to drive. Once started, the process becomes one of expanding accommodation to the automobile. New York's Mayor Wagner declares that older cities are approaching a point of saturation in the volume of passenger automobiles.[38] But the central areas of our older cities are not immutable for, as Mumford says, the automobile is the "sacred cow" of the "curious religion of technology" for which "no sacrifice in daily living, no extravagance of public expenditure, appears too great.[39]

The case of Los Angeles demonstrates that these areas will compromise themselves, though without particular efficacy, to the demands of automobile usage. The piecemeal capitulation of downtown Los Angeles involved, in the period 1931 to 1956, the clearance of 51.3 acres, 63.2 percent of which represented conversions of property formerly occupied by buildings to vehicular "open-air" parking,

while 21.9 percent of new construction during this time was space in parking structures.[40] It is evident from the downward trend of retailing activity that this adaptation has not been a viable solution but it is one that is and assuredly will be emulated in other areas with similar results unless the effects are imaginatively understood and planned beforehand.

It may be that the automobile, as we know it and for the purposes for which it is presently used, is already obsolete and that the present period of enormous effort to make over our physical environment in order to get along with it is merely an interlude, that the secularly increasing pressure of urban population and advancing technology will surely dictate new principles of urban organization for the future. In the meantime the delayed effects of the automobile revolution coming out of Los Angeles and pervading the rest of the country, including the problem of smog which is becoming increasingly ubiquitous, will continue to challenge the best efforts of planning. The Los Angeles region demonstrates great significance as the progenitor of urban change throughout the United States. If the tendencies manifest in the Los Angeles area were better understood, this would be tantamount to developing a science of urbanism and a substantive basis for public policies addressed to the improvement of the quality of life in the metropolitan areas which increasingly comprise the typical human ecological setting in this and other countries.

[37] T. R. Fyvel, "In Search of Urbanity," *Confluence*, Spring 1958, p. 51.

[38] *The New York Times*, January 12, 1959, p. 1.

[39] *Op. cit.*, p. 213.

[40] Henry A. Babcock, Consulting Engineers, *Report on the Economic Phases of the Bunker Hill Renewal Project, Los Angeles, for the Community Redevelopment Agency of the City of Los Angeles, May 15, 1956.* Moreover, such adaptation has been taking place at an increasing rate: 475 parking stalls were created in 1953, 794 in 1954, 859 in 1955, 1,302 in 1956, 1,339 in 1957, and 1,462 in 1958. (Figures compiled by the Master Plan Division of the City Planning Department, City of Los Angeles.)

Costs, Productivity and Welfare Problems of the Local Transit Industry

By EDWARD SUSSNA[*]

THE public utility concept derives from two basic premises: one, some industries are subject to such spectacular economies of large-scale production as to make their operation under monopolistic conditions socially desirable; two, this monopoly power must be publicly regulated in the interests of consumers so that the monopolist is limited to a fair return on his investment. The concept is basically a negative one—restricting an otherwise greedy monopolist from exploiting the public. It, in itself, offers no solution to the problems which arise in a dynamic economy. Public utility regulation was not designed to protect a utility from technologically superior competitors; the utility itself must adjust to the competition, or fail. Local public utilities, for example, have found their monopoly over urban transportation threatened by competition from automobiles to the point where even large increases in fares have not yielded reasonable returns on investment. The economist generally feels no special concern for firms or industries which are unable to survive changed consumer habits or competition from technologically superior enterprises. In fact, such attrition must be expected as a cost of economic progress. However, when a utility which is providing a presumably necessary service is threatened with extinction, some special public concern is clearly appropriate.

Study of the problems of the local transit industry emphasizes two major issues. First, do the pricing techniques presently used in the industry offer results which best further the public interest? Second, is regulation in its present form capable of providing desirable results? Data presented in this paper indicate that the local transit industry is a utility which in many cities has not provided a fair return on its investment nor kept the quantity and quality of its service up to the rate of growth of the cities it serves. Several important aspects of welfare economics can be used as a framework for suggesting policy solutions to the industry's financial dilemma. It will be shown that the public utility concept is basically a negative one and incapable of providing solutions to changing urban transportation requirements. Also, it will be suggested that consideration be given to the possibilities of substituting a multi-part fare structure for the flat fare system currently in use and of subsidizing a profitless transit company.

The Financial Record of the Local Transit Industry

Utilities, of course, differ from other privately-owned industries in that the former must obtain prior permission before altering its rate or service schedule. The ability of a utility to maintain a "fair" rate of return on its investment during an inflation depends upon its ability to pass along higher costs to the consumer in terms of higher rates and /or to make real (if not money) reductions in costs. A survey made by the author indicates that public utility commissions have generally permitted without much delay rate increases to compensate for higher costs.[1] Some utilities, electric

[*] Assistant Professor of Industry, University of Pittsburgh.

[1] Edward Sussna, *Public Policy Toward Labor-Management Relations in Local Public Utilities in Selected States, 1947-1952* (Unpublished doctoral dissertation, University of Illinois, 1954), pp. 55-56.

LAND ECONOMICS

and gas for example, still reasonably free of competition, have retained their customers and have been able to raise productivity to keep costs from rising less rapidly than they would otherwise in an inflation. This is not to infer that inflation is a cause of the financial crisis in the local transit industry. The real causes are changes in consumer habits and competition from lower-priced (and perhaps even under-priced) methods of transportation. It is more likely that inflation has veiled and perhaps even aided the industry in opposing real wage increases in face of a decline in real demand.

The local transit industry has found it difficult to raise productivity or pass on higher costs to its customers, even when commissions have permitted higher rates. Table I offers data on the financial status of the industry since 1945. It is clear that there has been a general deterioration in the economic condition of the transit industry. Operating revenue has been declining since 1953 in spite of the price inflation. Operating income fell sharply from $148,730,000 in 1945 to $55,690,000 in 1956. The plight

of the transit industry becomes even more apparent when the data on the number of revenue passengers and on revenue vehicle miles operated are examined. Thus, Table II depicts the drastic decline in revenue passengers from 18,981,900,000 in 1945 to 8,756,-000,000 in 1956, a decrease of 53.9 percent. This decline has been most serious in surface railways (from 7,080,-900,000 in 1945 to 625,000,000 in 1956) although this to a large degree reflects the shift from trolley cars running on fixed tracks to buses and trackless trolleys. But this slack was not taken up by the buses as evidenced by the decline in the number of revenue passengers they carried. Nor did subway and elevated lines take up the slack for they too experienced a decline from 2,555,100,000 in 1945 to 1,749,000,000 in 1956.

The narrowing scope of operations of the local transit industry is indicated by the data presented in Table III. The total of revenue vehicle miles declined substantially from 3,253,800,000 in 1945 to 2,366,600,000 in 1956, a decrease of 27.3 percent. Again the decline was not even for all modes of transit. Thus, sur-

TABLE I—RESULTS OF TRANSIT OPERATIONS IN THE UNITED STATES IN 1945, 1950-1956[1]

Year	Operating Revenue	Operating Expenses (including depreciation)	Net Revenue	All Taxes	Operating Income	Operating Income as percent of Operating Revenue
	(Thousands)	(Thousands)	(Thousands)	(Thousands)	(Thousands)	
1945......	$1,380,400	$1,067,140	$ 313,260	$ 164,530	$ 148,730	10.77
1950......	1,452,100	1,296,690	155,410	89,040	66,370	4.57
1951......	1,472,700	1,331,270	141,430	95,340	46,090	3.13
1952......	1,501,300	1,369,560	131,740	101,990	29,750	1.98
1953......	1,513,100	1,370,700	142,400	97,350	45,050	2.98
1954......	1,471,800	1,337,260	134,540	89,700	44,840	3.05
1955......	1,426,400	1,277,370	149,030	93,320	55,710	3.91
1956......	1,416,100	1,271,480	144,620	88,930	55,690	3.93

[1] American Transit Association, *Transit Fact Book*, 1957, p. 4.

even
the
on
ex-
the
gers
56,-
53.9
nost
80,-
)56)
ects
on
eys.
the
the
car-
ted
ex-
000

s of
by
The
ned
945
of
not
ur-

ome
f

TABLE II—Revenue Passengers Carried on Transit Lines of the United States Distributed by Types of Service in 1945, 1950-56[1]

Calendar Year	Railway			Trolley Coach	Motor Bus	Grand Total	Percent Change From 1945
	Surface	Subway and Elevated	Total				
	(Millions)	(Millions)	(Millions)	(Millions)	(Millions)	(Millions)	
1945............	7,080.9	2,555.1	9,636.0	1,001.2	8,344.7	18,981.9	
1950............	2,790.0	2,113.0	4,903.0	1,261.0	7,681.0	13,845.0	−27.1
1951............	2,171.0	2,041.0	4,212.0	1,231.0	7,438.0	12,881.0	−32.1
1952............	1,714.0	1,982.0	3,696.0	1,201.0	7,125.0	12,022.0	−36.7
1953............	1,403.0	1,903.0	3,306.0	1,137.0	6,593.0	11,036.0	−41.9
1954............	1,039.0	1,781.0	2,820.0	993.0	6,045.0	9,858.0	−48.1
1955............	845.0	1,741.0	2,586.0	869.0	5,734.0	9,189.0	−51.6
1956............	625.0	1,749.0	2,374.0	814.0	5,568.0	8,756.0	−53.9

[1] American Transit Association, *Transit Fact Book*, 1957, p. 8.

face railways declined most (again largely as a result of the shift to buses and trackless trolleys) but subway and elevated, trolley coaches, and motor buses all showed some degree of decline.

Having presented these facts, the problem is to explain the decline of the local transit industry (a public utility and therefore presumably of vital importance to the public) in a period of general prosperity. The next section discusses some of the factors which account for this anomaly.

Factors Contributing to the Decline of the Industry

A Decline in Passenger Volume. Quite obviously, the significant decline in pas-

TABLE III—Revenue Vehicle Miles Operated in the United States by Each Type of Transit Vehicle in 1945, 1950-1956[1]

Calendar Year	Railway			Trolley Coach	Motor Bus	Grand Total	Percent change from 1945
	Surface	Subway and Elevated	Total				
	(Millions)	(Millions)	(Millions)	(Millions)	(Millions)	(Millions)	
1945............	939.8	458.4	1,398.2	133.3	1,722.3	3,253.8	
1950............	463.1	443.4	906.5	205.7	1,895.4	3,007.6	− 7.6
1951............	387.6	424.0	811.6	208.8	1,893.0	2,913.4	−10.5
1952............	321.2	400.4	721.6	215.2	1,877.7	2,814.5	−13.5
1953............	273.7	391.1	664.8	211.7	1,819.0	2,695.5	−17.2
1954............	215.1	376.3	591.4	196.7	1,760.7	2,548.8	−21.7
1955............	178.3	382.8	561.1	176.5	1,709.9	2,447.5	−24.8
1956............	132.9	387.1	520.0	165.7	1,680.9	2,366.6	−27.5

[1] American Transit Association *Transit Fact Book*, 1957, p. 11.

senger volume, as depicted in Table II, has been a major cause of the financial plight of the industry. The problem is now to determine why there has been this decline. Several possibilities arise. One, increases in fares have discouraged transit use. Two, cheaper or more convenient modes of transportation have become available. Three, the increases in real income enjoyed by Americans since 1945 have led them to regard local transit travel as an "inferior" service, causing a change in consumer habits. Four, radical changes in urban population distribution have complicated transit investment decisions.

With regard to transit fares, evidence compiled by the United States Bureau of Labor Statistics indicates that the price index for public transportation, which consists mainly of local transit fares, rose from 117.9 in June 1950 to 180.2 in July 1957 (1947-49 = 100) while the indexes for private transportation and for all items rose from 106.6 to 125.6 and 101.8 to 120.8, respectively, during this period.[2] Does this mean that local transit is pricing itself out of business when it should instead be holding fares stable or perhaps even cutting them in an effort to stimulate transit use? It appears that, for some transit lines, offering the type of service they now do, there is no point on relevant portions of their demand schedules which permits a reasonable return on their investments.

Second, the attainment by even low-income families of the economic wherewithal to operate automobiles has offered serious competition to transit use. In 1957, cars were owned by 72 percent of all consumer units (with higher percentages in the urban areas) and this percentage can be expected to increase with increases in real national income.[3]

[2] United States Department of Labor, Bureau of Labor Statistics, Mimeographed release, August 22, 1957.

[3] Automobile Manufacturers Association, *Automobile Facts and Figures* (Detroit, Michigan, 1957) p. 32.

If parking sites are readily available and car pooling is used, direct costs of automobile travel are likely to be less for a group using one vehicle than if each person were to use public transportation. It is true that direct (or variable) costs are not a good measure of the cost of automobile travel. Many fixed items as depreciation, insurance, maintenance, etc., should also be included. But, in the short run and after the decision to own a car is made, variable costs certainly become more relevant. There is also some reason to question whether the social costs (in terms of adding to already crowded streets) are not significantly higher than the private (fixed and variable) costs borne by the automobile user. Despite the facts that automobile use has enlarged the problems of urban mass transportation and that the economic vitality of many commercial and industrial areas is being sapped thereby, there has been reluctance on the part of planning officials to take appropriate, if drastic, action.

Third, for a large number of people the choice between using public or private transportation does not depend only on simple cost comparisons. With the general increase in real incomes, significant numbers of people have come to prefer to use their automobiles even when the direct costs thus incurred exceed transit fares. And perhaps this decision is not as uneconomic as it appears because the services of private and public transportation are not strictly comparable. The former is less time-consuming, offers door-to-door service, and is more comfortable. Local transit has in some cases been upgraded to the point where it attracts upper-income riders but has not been generally true.

Fourth, the past twenty years have seen remarkable changes in urban popu-

lation size and dispersion. With the growth of suburbs the traditional concept of the city has given way to that of the metropolitan area. The population of the latter has grown greatly because of the movement to the suburbs but the population densities of the central areas of many cities have either stayed constant or actually declined. Fixed transit systems have therefore suffered because they serve less populated areas and cannot readily shift their facilities to accommodate the new population trends. Even trackless systems find the new distribution unfavorable because suburbanites are so widely dispersed through the metropolitan area as to make service to outlying points unprofitable.

B. *Technology and Operating Expenses.* Faced as they are with unfavorable demand conditions, local transit companies must reduce operating expenses if they are to remain solvent. The largest item included in operating expenses is labor costs, accounting usually for 60 to 70 percent of the total budget. Between 1947 and 1955, weekly and hourly wages of transit employees rose from $57.14 and $1.22 to $80.60 and $1.87, respectively. These are increases of 41 and 53 percent. For the same period, wages in the manufacturing industries rose from $43.74 to $76.52 weekly, and $1.08 to $1.88 hourly, increases of 75 and 74 percent respectively.[4] There is no reason to believe, on the basis of this evidence, that the plight of the transit industry can be attributed to its paying wages higher than those prevailing in other industries. If anything, the transit industry has bargained harder with its employees in efforts to keep costs down than other utilities have with their employees. Thus, the strike record of the local transit industry is worse

than that of local electric utilities and the average of all industries. Data provided by the United States Bureau of Labor Statistics show that in 1954 man-days idle as a percentage of total estimated working time was .64 for transit workers, and only .005 and .21 for employees of electric utilities and all industries, respectively. To expect the transit industry to reduce its costs by depressing wages paid its employee is neither practical nor desirable. If wages are kept significantly lower in local transit than in comparable industries, difficulty would be met in attracting an adequate labor force, in quality and quantity. Even if such low wages were possible, there would be serious questions as to the propriety of subsidizing and thus artificially prolonging a basically inefficient industry.

A more practicable approach to the problem of reducing operating expenses is to increase the use of capital equipment to raise employee productivity. This would not reduce the pay of individual workers but would reduce the total wage bill by substituting capital equipment for labor. This, of course, is the way in which most industries have raised their productivity. The transit industry has resorted to such practices as eliminating the second man (conductor) on street cars, reducing service during slack periods, or permitting general deterioration in service and maintenance in efforts to increase productivity. Such efforts offer limited relief only and have not really solved the major problem of increasing the speed at which riders can be transported through crowded city streets. To achieve this goal would require not only significant technolgical advances but also considerable help from municipal officials in reducing traffic congestion.

[4] Data are from the regular series of United States Bureau of Labor Statistics, *Monthly Labor Review.*

Statistics provided by the American Transit Association depict the decline in physical productivity of the local transit employee. Although the number of employees declined from 242,000 to 186,000 between 1945 and 1956, the number of revenue vehicle miles (output in this case) declined even more drastically during this period—from 3,254,-000,000 to 2,367,000,000. In productivity terms, as measured by output per man, the decline was from 13,450 in 1945 to 12,730 in 1956—a decrease of five percent. More significantly, the decline was probably greater in the big cities where the mass transportation problems are more acute. It must be noted that revenue vehicle miles per employee is at best a rough measure of productivity. This concept, for example, ignores such basic changes as those in the input mix (changes in the relative use of the factors of production), number of hours worked per year and intensity of that work, quality of service offered, number of passengers carried, etc. It is also important to remember that these figures are of physical, not value productivity. Despite these limitations, the statistics offer some insight into the problems of local transit.

The key to increased productivity is in getting transit systems out of competition for space on crowded streets. To do this would involve a radical departure from the form of transit service provided in most urban areas at present. To reduce the competition for limited street space would entail three major alternatives. One, the city could provide the transit company with a private right-of-way on street level. This might mean that certain streets, probably key ones, would be barred to private transportation, or at least that lanes in these streets would be reserved for public carriers. A program like this, however, might be politically if not legally unfeasible. A second alternative would be to get the transit system above street level. But this would have to be done in a manner somewhat different from the way in which the old elevated lines were constructed if public accceptance is to be expected. The din and gross unsightliness of the elevated system would have to be eliminated. A possible solution would be the monorail—a system already in use in some European countries. A third alternative would be increased use of subway rapid transit lines. Subways offer the distinct advantages of speed, reliability under all sorts of weather conditions, and non-competition with street traffic. However, both subway and elevated systems suffer severe limitations in that they require large initial investments. Considering the financial plight of most large transit systems there is no reason to expect this capital to be supplied by private investors. Furthermore, because of the heavy fixed investment only the largest cities could provide sufficient passenger volume to justify the new construction. Finally, subway and elevated lines are inflexible and cannot be easily modified to accommodate changing residential patterns.

C. *Peak Load Problems and the Fare Structure.* Transit systems must have enough equipment available to meet not average demand but peak demand for its services. The closer that peak demand is to average demand the better will be the use that the company will make of its equipment and manpower. But, accentuating the problems of the transit industry has been the widening gap between average and peak passenger traffic. The peaking problem has been of two types; one, relatively big decreases in weekend travel; and two, peaking of weekday travel. This pattern is demon-

strated in statistics furnished by the New York City Transit System. As depicted in Table IV, Saturday, Sunday, and

TABLE IV—Average Number of Saturday, Sunday, and Holiday Passengers as a Percentage of Average Number of Weekday Passengers, New York City Transit System, 1948-1956[1]

Fiscal Year Ended June 30	Rapid Transit Lines	Surface Lines	System Total
1948.........	55.03%	79.67%	60.86%
1949.........	50.54	74.12	57.30
1950.........	47.70	67.41	53.34
1951.........	46.16	63.92	51.12
1952.........	45.47	61.22	49.50
1953.........	42.80	59.79	47.05
1954.........	40.66	56.30	44.27
1955.........	38.86	53.72	42.20
1956.........	38.64	51.85	41.60

New York City Transit System, *Annual Report for the Year Ended June 30, 1956.*

holiday traffic, as a percentage of weekday traffic, has fallen off substantially. Between 1948 and 1956, the percentage for the whole system declined from 60.86 to 41.60 percent; for rapid transit lines from 55.03 to 38.64 percent; and for surface lines from 79.67 to 51.85 percent. The weekday peaking trend is depicted in Table V. From these data it is evident

TABLE V—Peak Hour Rapid Transit Passengers as a Percentage of 24 Hour Total on a Typical Week Day, New York City, 1948, 1951, and 1956[1]

Fiscal Year Ended June 30	Two Morning Peak Hours 7-9 AM	Three Evening Peak Hours 4-7 PM
1948............	22.95%	29.35%
1951............	24.42	30.43
1956............	27.11	31.75

[1] New York Transit System, *Annual Report for the Year Ended June 30, 1956,* p. 17.

that in 1956 almost 60 percent of weekday riding was accomplished in five hours (7-9 A.M. and 4-7 p.m.) out of a possible 24. This heightened peaking trend has serious economic implications,

for the transit industry. It means that much equipment will be unused or underused and that large fixed costs will be spread among fewer passengers. Such costs as electric power, depreciation, maintenance, and inspection are almost entirely unrelated to the volume of traffic. Many labor costs (e.g., attendants at subway booths, operating personnel) are variable only within relatively narrow limits. This implies little flexibility in adjusting operating costs to changing riding habits.

If costs cannot be adjusted in response to the increase in peaking a more practicable approach might be to engage in some form of economically-justifiable fare discrimination. Fare discrimination is not new to the transit industry. Fare reductions have been made available on a limited basis to off-peak passengers in some cities. A more common type of discrimination is that of the reduced fares generally made available to students. These reductions, however, are not consistent with economic principles for the students often travel during rush hours. Since they are partially responsible for the cost of additional equipment necessary to meet increased peak needs there is no economic rationale for giving them lower fares. There may, however, be important political or welfare reasons for the discrimination. The traditional economic requirements for the success of discrimination would be present in the transit industry for riders can be grouped into identifiable classes with distinctly different demand functions. Few rush-hour passengers are able to shift their transit use to an off-peak (lower-fare) hour. Finally, it is impossible for the service to be re-sold between passenger groups. Fare discrimination, to the extent that it encourages more off-peak travel, would increase total revenue more than it would operating costs as it would

involve better use of equipment and manpower already employed to handle rush-hour traffic.[5] But off-peak demand is probably not as elastic as might superficially appear for the choice between private and public transportation is not always based on economic cost comparisons. Thus, women shoppers may prefer the convenience of automobiles to the transit lines, cost considerations notwithstanding.

The fare structure should also provide for differences in length of trip. A New York subway rider pays the same fare for a trip of ten blocks as he does for one of thirty miles. Fare differentials, broadly related to trip distance, would reflect costs more accurately. They would also serve to regain some of the passengers who have switched to other modes of tranportation or to walking in preference to paying full flat fare increases for short trips.

The objections to multi-part fare structures are both economic and political. Fare collections under such a system become more complex and hence more expensive. Also, certain types of discrimination which would logically reflect cost differences might be politically unpalatable. Thus, increased fares to rush-hour users, though justified on cost grounds, might have adverse political repercussions. It would seem that the major barrier to a multi-fare system is political; explanation to the public of the reasons for such a fare structure might lessen opposition to it.

Welfare Implications

It has been suggested that a multi-part fare structure might stimulate transit passenger demand and would reflect cost differences more equitably than does a flat fare. Possible modification of the tranditional profit and loss criterion for pricing should also be considered. Are there any reasons why the *whole* fare structure should be lowered below average cost level? On the basis of the welfare economics principles enumerated in this section thought should be given to the possibility of subsidizing the local transit industry under certain conditions.

As described by Scitovsky, "Welfare economics is that part of the general body of economic theory which is concerned primarily with policy."[6] He goes on to point out that welfare economics involves making value judgments, an activity which makes the scientific objectivity presumably achieved by the natural scientist unavailable to the social scientist, i.e., the economist. But of course value decisions continue to be made by public officials regardless of whether the economist feels that it is within his province to participate in these activities or not. Thus the economist has come to realize that most of his intellectual activity is meaningless unless it is somehow related to the solution of normative problems.

It would seem that the application of welfare economics criteria to a public utility is most compelling for, unlike other privately-operated sectors of our economy, utility managements and consumers are vitally affected by decisions of public officials. This regulation is inherent in the public utility concept and continues with or without advice of economists. It seems that economists would be derelict in their responsibilities if they did not contribute their knowledge to the decision-making activities of an industry in which decisions with profound welfare implications were already being made. These publicly-made, or at least publicly-regulated, decisions have to do

[5] A more complete discussion of transit fare discrimination is found in William S. Vickrey, *The Revision of the Rapid Transit Fare Structure of the City of New York*, Mayor's Committee on Management Survey of the City of New York, Technical Monograph 3, 1952.

[6] Tibor Scitovsky, "The State of Welfare Economics," *American Economic Review*, June 1951, p. 303.

with such items as rates, quality and extent of service, investment, etc.

Three aspects of welfare economics, not unrelated, are significant to policy-setting in the local transit industry. The first involves the problem of economic efficiency—optimum allocation of resources. The second involves equity considerations —optimum income distribution. As will presently be observed, the second is much more complex (and hence less likely to be resolved) than is the first. The third relates to possible divergences of private and social costs.

The first consideration is that of relating local transit operation to economic efficiency. This involves inquiry into whether, on welfare grounds, public policy should be oriented towards sustaining a privately-operated enterprise. Implicit in the discussion, until now, has been the assumption that the transit industry is of vital importance to the city and that all reasonable efforts must be made to sustain it. But, some may argue, why bother sustaining a private industry which is so obviously in financial peril? Should not the transit industry be capable of financial success without subsidies or are there special criteria we ought to apply to the industry which entitles it to public help? We have got to remember that the local transit industry is a publicly-regulated (and in several cities, a publicly-owned) monopoly. Five of the transit systems of the ten largest American cities are publicly owned. These are New York, Chicago, Detroit, Cleveland, and Boston. Traditional concepts under which the profit and loss calculus serve to allocate resources optimally need not apply in this case. It can be theoretically demonstrated that a policy of charging less than average cost (in a decreasing cost industry) and making up the deficit through subsidies is consistent with maximizing both public

welfare and consumer sovereignty. Once we have substituted commission regulation or public ownership for competitive pricing we might just as well be prepared to modify traditional pricing practices to conform logically with the new framework.

The case for charging less than average cost is based on the general welfare principle that optimum resource allocation is obtained when production is extended to a point where price equals marginal cost. Under competitive conditions equilibrium will be achieved when price, average cost, and marginal cost are equal, thus covering all of the firm's costs. For a monopolistic utility, however, marginal cost and price both may lie and intersect beneath average cost for the most relevant portions of output because of indivisibilities and high fixed costs in the short run, and increasing returns to scale in the long run. There is good reason to believe that the transit industry is subject to decreasing costs as are most public utilities. Such a tendency is rather easily provable for rapid transit systems not in competition with street traffic. But the possibility of being subject to increasing returns exists for even buses and street cars. A bus company, for example, can continue to enjoy increasing returns for significant expansions in service. Whether bus size is increased or service frequency is increased (thus upgrading service quality), average costs will decline. The only limitation to prospects for declining costs is that the buses *themselves* are so plentiful as to get in each other's way. That public and private transportation must often compete for limited street space is not relevant in determining whether either is subject to decreasing costs. The competition for crowded streets is a problem best handled in discussing the

divergence between private and social costs, taken up later in this section.

Once this decreasing cost assumption has been made, it is quite economically sound to support operation of such an industry even when its demand schedule always lies below its average cost curve. Following Marshall, and more recently Hotelling, output of such an industry should be expanded until marginal costs and price are equal.[7] The resulting deficits would be met by subsidies raised by taxation which itself would not interfere with use of the subsidized activity.

The second welfare aspect of the local transit situation is that concerning the distribution of income among transit service users. Welfare considerations require that income satisfaction for the group be maximized by distribution of income in such a way that equal marginal expenditures of all persons in the group provide equal marginal satisfaction. Expenditures can be measured in monetary terms. Interpersonal comparisons of satisfactions, however, defy objective measurement. This latter limitation has not prevented public officials from promulgating policies which allegedly correct maldistributions of income. These redistribution techniques are rough and difficult to justify on purely objective grounds but have been generally accepted (if for no other reason than that there are many more people who expect to gain than to lose from such redistribution).

The third aspect of welfare economics applicable to the local transit industry is that concerning the divergence between private and social costs. A prominent argument advanced by public transpor-

tation groups is that private transportation is being subsidized in that general tax revenues are being used to improve city streets. It seems more likely, however, that the costs of streets are more than met by taxes on urban automobile owners and that in fact urban taxpayers are subsidizing rural road users. But cost considerations should not be limited to money terms only; they should include delay and risk which are important parts of the transportation service. On a crowded street each additional automobile user inflicts a cost (money, delay, and risk) on public and other private transportation which is less than he himself bears. Thus social costs exceed his private costs. And it is likely that the divergence of social and private costs is greater for private than for public transportation. Or, to put it differently, the local transit industry bears a significant part of the social cost which is generated by private transportation. The infliction of this additional cost on the transit industry is largely responsible for its inability to benefit from the increasing returns it might otherwise enjoy.

The welfare aspects of the fare and subsidy structure of the transit industry might appear less important than welfare considerations on expenditures which take numerically larger percentages of family incomes. But the transit problem has received a more than proportionate amount of public concern for the following two reasons. First, the public is aware that control is possible over transit fares but not over price practices of other private industries. Second, the lower income families are more dependent upon public transportation than are the higher income families.[8] A combination of these factors emphasizes to officials in

[7] Alfred Marshall, in his *Principles of Economics*, 8th ed. (London, England: Macmillan and Co.), pp. 472-473, quite clearly recognized this special case but it remained for Harold Hotelling to develop it to a logical conclusion in "The General Welfare in Relation to Problems of Taxation and of Railway and Utility Rates," *Econometrica*, July 1938, pp. 242-269.

[8] In 1957, 61 percent of the lowest and 32 percent of the next to lowest family income quintiles did not own an automobile. Data are from Automobile Manufacturers Association, *op. cit.*, p. 33.

many cities that it is a political necessity to keep fares down. But, even apart from political considerations, a good case can be made for establishing a fare structure which provides less revenue than is necessary to cover total costs. Assuming that government subsidy or ownership become necessary, increases in fares have two undesirable results. They are especially burdensome to low-income groups who often must use public transportation and they discourage fuller (and hence more nearly optimum) use of the transit system.

Admittedly, there is some disagreement as to theoretical validity of welfare economics and much cynicism as to its practical applications.[9] Marginal cost pricing is not always practicable nor even desirable in that it requires that subjective decisions be made on ordinal utility functions. However, these limitations need not preclude the use of broad and reasonably uncomplicated fare discrimination policies which serve to enhance economic welfare.

Conclusions

In some ways discussion of the economic problems of the local transit industries is analagous to talk about the weather—there is much talk but very little done about it. Almost everyone agrees that public transportation systems are indispensable to solving urban traffic problems. Emotions run high in attacking transit management and labor as irresponsible and unmindful of the public interest when strikes occur in that industry. Why is it that the financial position of the local transit industry is deteriorating in spite of all the lip service paid to its public indispensability?

That the financial status of the local transit industry is deteriorating is em-

[9] The most comprehensive analysis of the limitations of welfare economics is contained in I. M. D. Little, *A Critique of Welfare Economics* (2nd ed.) (Oxford, England: Oxford University Press, 1957).

phasized by the data presented in Tables I, II, and III. Since 1945 revenue passengers carried and revenue vehicle miles operated have declined substantially. Operating income as a percentage of operating revenue has also fallen drastically in that period. Qualitative degeneration, as relating to frequency of service and passenger comfort, has accompanied (and perhaps added to) the financial crisis. The major causes of this plight are these: (1) On the supply side, real costs per passenger carried have increased. Where many industries have partially at least met rising costs by more intensive use of capital equipment, the transit industry, highly labor-intensive, has until now made no significant use of labor-saving devices. (2) On the demand side, passenger volume has been less inelastic in response to fare increases than has demand for the services of other utilities. The big gap between social and private costs of urban automobile travel, wide ownership of automobiles, and the attitude of many people that public transportation is essentially an "inferior" service in a period of rising real incomes all have contributed to the industry's problems. (3) The riding habits of transit passengers cause more peak load concentration than ever existed previously. While passenger traffic for all hours and days has generally fallen off the decline has been relatively greater in weekend and weekday non-rush hour travel. Thus, heavy investments in equipment must be made to satisfy the peak needs of rush hour passengers—usually about five hours per day for five days per week. Consequently, this large fixed plant (and usually manpower also) is unused during most of the week.

If the local transit industry is to be salvaged, and the public and traffic experts feel that it must even though they seem unwilling to pay the necessary fares,

several important lines of action are indicated. These are: First, in the short run, where equipment is relatively fixed, transit companies could pursue policies which encourage better use of their equipment and manpower. Such practices could include fare discrimination, express service on heavily-travelled routes, and off-peak rental of equipment for such purposes as sightseeing, group charter trips, etc. Second, in the long run, where major investment decisions will be made, thought must be given to transportation modes which are significantly more efficient than present one. Substantial increases in productivity can come about only when speed and service are improved. Such improvements, in turn, will come only when the public transportation systems are taken out of competition with other modes of travel on streets already physically overcrowded. Elimination of such competition would involve the city's granting private rights-of-way on street level or more realistically, the use of elevated or subway systems. Whether private capital would be willing to invest as heavily as needed for these new systems is problematical, in view of the poor return paid on existing transit investment. Third, and closely related to the first two proposals, is the proposition that the transit industry's problems are insoluble without positive public policy. Public officials will first have to recognize that, for the larger cities at least, private transportation is

inadequate in handling mass transportation needs. Moreover, once the need for public transportation is generally accepted as being consistent with optimum resource allocation, cities may have to offer subsidies to private systems to keep them financially solvent. Or it might be necessary for a city to assume management of the system, where private owners are reluctant to continue operation or where they are obviously shirking their public responsibility while accepting public subsidies.

This discussion would not be complete without mention of the relationship between the local transit problem and the public utility concept. Even if it is conceded that the transit industry is in financial straits largely through no fault of its own the question arises as to whether the concept really protects the public as it was intended. As long as the private owners were able to secure adequate profits, private monopoly power was regulated (and protected) and extolled as being consistent with a private enterprise system. But regulation at best is a negative weapon protecting the public from the most serious excesses of monopoly. Regulatory bodies *cannot require* that socially desirable decisions be made by the monopolist. Thus, the public is faced with a situation in which it must assume the operation and tax burden for a utility only after the utility has ceased to be profitable.

Corporate Tax Allocation in Wisconsin

By JOHN A. WILKIE*

RECENT decisions of the United States Supreme Court have re-emphasized the need for states to re-examine the statutes by which they allocate interstate business and income for tax purposes. The character of these statutes can substantially affect the conditions of competition and the growth of cosmopolitan business.

This article is an examination of the allocation procedure of one state—Wisconsin. It is a case-type study. The author hopes it may find wider application than in the one state here directly analyzed.

The Problem

A survey of 1949 Wisconsin corporate net income taxes shows that, of 10,309 taxpaying corporations 9,401, or 91.2 percent, had 100 percent of their income attributable to Wisconsin (as determined under Wisconsin law). Those corporations, however, had a net taxable income of 245.6 million dollars or only 47.2 percent of the total net taxable income and only a slightly smaller percent of the total tax. Thus 908 or only 8.8 percent of all taxpaying corporations had less than 100 percent of their income attributable to Wisconsin (as determined under Wisconsin law) but they accounted for 275.3 million dollars or 52.8 percent of the total net taxable income and a slightly greater percentage of the total tax.[1] From this it is apparent that a large share of Wisconsin's total net taxable income and corporate income

tax is derived from corporations with income attributed to more than one state. Thus the problem of allocating the income of a corporation which is derived from sources or attributed to sources in two or more states is an important one in Wisconsin.

Source of Income Jurisdiction

Wisconsin levies its income tax directly "upon such income as is derived from property located or business transacted within the state."[2] Also, with respect to situs of income, "persons *engaged in business* within and without the state shall be taxed only on such income as from *business transacted* and property located within the state."[3] Also among gross income inclusions is listed "all profits derived from the transaction of business or from the sale or other disposition of real estate or other capital assets, etc."[4] The administrative rule relating to this section provides: "All of the income from *business carried* on in Wisconsin is taxable. The fact that a person or corporation is licensed to do business in Wisconsin is evidence that it is doing business in the state within the meaning of this chapter. However, a person or corporation may be doing business in this state within the meaning of the chapter even though not licensed."[5] Where all income of a mercantile or manufacturing business is derived from *business transacted* and property located within the state then it is not considered as being engaged in business within and without the state and its in-

* Research Associate, Wisconsin Tax Impact Study.

[1] One hundred and twenty-eight of these "multistate" corporations with less than 100 percent of their income attributable to Wisconsin each had a total Wisconsin net corporate income tax liability of over $20,000 and accounted for 252.9 million dollars net taxable income while the remaining 780 such corporations account for only 22.4 million dollars net taxable income.

[2] Section 71.01 (1) Wisconsin Statutes 1957.

[3] Section 71.07 (2) Wisconsin Statutes 1957.

[4] Section 71.05 (1) (g) Wisconsin Statutes 1957.

[5] Wisconsin Administrative Codes, Tax 2.69, *Income from Wisconsin Business.* It also adds that "in all cases of doubt, the complete facts should be reported to the tax commissioner or assessor of incomes for determination."

come follows the situs of the business from which it is derived.[6]

Under these provisions there have been certain court rulings which show their restrictiveness. For example, Wisconsin has held that the carrying on of various sales activities outside the state does not mean that the corporation is "doing business" outside the state for apportionment purposes.[7] Sales agents of the corporation had no authority to bind the corporation in any contract. However, the corporation through its agents performed certain acts in other states. The name of the company itself appeared not only on all the transactions for its account but upon the entrance of the agent's office and in the telephone and other directories of the city. Such acts were clearly for the purpose of promoting the products of the corporation and should be sufficient to provide a basis for direct income tax jurisdiction.

In another case a domestic corporation had its principal offices and manufacturing plant in Wisconsin. All orders received were accepted or rejected at its principal office in Appleton. The company, however, maintained an office in Chicago with one sales employee as well as some clerical help and this office was used by the president of the company while in Chicago interviewing the trade and doing other business. Office furniture and equipment were located there; the corporation was listed in the telephone directory in Chicago; and there was some correspondence between the Chicago office and the trade. No stocks of goods were maintained outside Wisconsin. The Tax Commission held that all of the corporation's income was derived from business transacted and property located within the state and that the

taxpayer was not entitled to any apportionment.[8] It is submitted, however, that the corporation did own and used capital in Illinois, employed labor there, and did other acts there for the purpose of ultimate profit and that such activities should be sufficient to provide a basis for direct income tax jurisdiction.

In another case involving a construction job where a large part of the actual labor and materials for the job was provided or produced by the buyer the contractor was not held to be "transacting business within Wisconsin" even though it was responsible for the entire job and performed constant technical supervision during the process of construction.[9] Clearly these supervisors (employees of a foreign corporation) were performing services within the state that were in connection with the income-producing function of construction engaged in for pecuniary profit. Where a foreign corporation upon assuming a contract to erect a building in Wisconsin sublets all the work to subcontractors and puts supervisory men on the job to see that the best construction practices are applied, that corporation through its employees is performing acts in Wisconsin in connection with an income-producing function of the corporation for the purpose of financial profit. Income derived from or attributable to such activities should be taxable by the state where they are performed.[10]

It may be concluded from the above review of a few of these cases relating to "business income" not being subject to apportionment in Wisconsin that a change is needed in the statute to pro-

---

[6] Based on Section 71.07 (1) and (2) Wisconsin Statutes 1957.

[7] See *Trane Co.* v. *Wisconsin Tax Commission* (1940) 235 Wis. 516, 292, N,W. 897.

[8] Appeal of Combined Locks Paper Co., Wisconsin Tax Commission Decision, 1939.

[9] *Koppers Co.* v. *City of Milwaukee*, 191 Wis. 397, 211 N.W. 147.

[10] For other cases see *Aluminum Goods Mfg. Co.* v. *Wisconsin*, Circuit Court, Dane County, Wisconsin, February 6, 1941 and American Stores Dairy Co. v. Wisconsin, 246 Wis. 396, 17 N.W. (2d) 596.

vide for a broader concept of what is a "source" of income. There are at least three possible alternatives that might be adopted. One is the "source-of-income" method, an approach which no state has as yet adopted. The second, the "activities of the corporation or its representative in whatever capacity" method is similar to that used by Georgia. The third is the "property sited and activities carried on" method used by California. These methods are defined and compared elsewhere by the author.[11]

Determination of Domicile

Wisconsin law provides that "all other income, including royalties from patents, income derived from personal services, professions and vocations and farm land contracts, mortgages, stocks, bonds, and securities or from the sale of similar intangible personal property shall follow the residence of the recipient except as provided in S. 71.08."[12] For corporations, residence is construed in practice to mean the legal domicile. It is submitted that in the case of corporations the tax situs of these various types of income should be the commercial domicile of the corporation. Commercial domicile should be defined for income tax purposes on the basis of where the greatest benefits are received and greatest opportunities are afforded: that is, in the state of the primary locus of its principal business activities rather than exclusively in the state where a corporation was incorporated or where the "principal office" is located. The latter does not in many cases correspond to the state of greatest benefit and opportunity. As a definite and reasonable means of determining the state of primary locus of a corporation's principal business it is recommended that an activity ratio com-

posed of the two factors of tangible property and labor employed be authorized except in those cases where tangible property is not a significant factor, in which case the labor factor alone would be used.[13] For corporations, such as airlines, trucking companies and railroads, the ratio should be derived from the special apportionment formulas which have been developed for these businesses.

Until legal domicile is eliminated as a basis for allocating non-apportionable income it may be necessary in order to avoid double taxation, to tax only fifty

[13] In order to compare the effect of using "legal domicile" and "principal offices of the corporation" as a basis for tax jurisdiction, a tabulation was made of the status of 133 corporations which paid over $20,000 in corporation income tax in Wisconsin for the fiscal year ending June 30, 1950, and which derived income from sources in several jurisdictions. Fifty-three of these, or about 40 percent of the total, were found to have been incorporated in Wisconsin, whereas 75, or 56.4 percent of the total, would have had a Wisconsin location on the basis of the "principal office of the corporation" criterion.

If taxing jurisdiction is defined as the place where the corporation carries on most of its operations, a further increase in the number of firms with a Wisconsin situs is indicated. Seventy-six, or 57.1 percent of the 133 corporations, reported (on the basis of the Wisconsin apportionment formula or separate accounting) over half of their operations attributable to Wisconsin, and 82 firms had over 40 percent of their operations in the state. It should be noted, of course, that a firm can have considerably less than 40 percent of its operations in a state and still have more of its operations in that jurisdiction than in any other state. Lack of data for other states made a tabulation on this basis impossible.

The Wisconsin apportionment ratio used in these calculations employs three factors: tangible property owned and used, cost of manufacturing, and sales at the office by or through which sales are accepted. The cost-of-manufacturing figure applies only to manufacturing concerns. Eliminating the sales factor (as would be the case under primary locus as defined in the text above) would cause little change in the number of firms with most of their operations in Wisconsin. On this basis 74 corporations have half of their operations in Wisconsin and 80 have over 40 percent of their operations in the state. Replacing the "cost-of-manufacturing" factor with a "payroll" factor would also result in little change.

For corporations smaller than the large corporations as defined above, of 730 firms studied only 81 were incorporated in Wisconsin, while 87 had their principal office in this state and 124 conducted over 40 percent of their operations in Wisconsin. This would indicate that large taxpaying corporations are more apt to be incorporated in jurisdictions away from their principal offices than are the smaller firms. The primary overall conclusion to be drawn, however, is that a substantial increase in the number of corporations domiciled within Wisconsin's taxing jurisdiction would result if domicile were defined on the basis of the locus of its facilities and operations (capital and labor) rather than on the basis of incorporation.

[11] See John A. Wilkie, "A Basis for Taxing Corporate Net Income," Taxes, September 1958.
[12] Section 71.07 (1) Wisconsin Statutes 1957.

percent of such income at its commercial domicile (as herein defined) with the other 50 percent still being taxed by the state of legal domicile.   On the other hand, some double taxation in this area might be tolerated to discourage incorporation in states other than that of "commercial domicile" as uniformly defined or perhaps to encourage a move to have federal incorporation of corporations which derive income from sources within more than one state.[14]

Allocation Method for Unitary Business

Wisconsin law provides that "the amount of such income attributable to Wisconsin may be determined by an allocation and separate accounting thereof, when the business of such person within the state is not an integral part of a unitary business *provided however that the department of taxation may permit an allocation and separate accounting in any case in which it is satisfied that the use of such method will properly reflect the income taxable by this state.*"[15]   It is recommended that the proviso clause be eliminated so that no separate accounting would be permissible in the case of a unitary business.[16] This is based upon the fundamental conclusion that separate accounting is appropriate only for separate business and is not appropriate for a unitary business operation.[17]   Of course, a corporation may have both a separate and unitary business in which case the separate

business would be separately accounted for.

Elsewhere in the statutes Wisconsin law provides:

"When any corporation liable to taxation under this act conducts its business in such a manner as either directly or indirectly to benefit the members or stockholders thereof or any person interested in such business, by selling its products or the goods or commodities in which it deals at less than the fair price which might be obtained therefor, or where a corporation a substantial portion of whose capital stock is owned either directly or indirectly by another corporation acquires or disposes of the products of the corporation so owning a substantial portion of its stock in such a manner as to create a loss or improper net income, the department may determine the amount of taxable income to such corporation for the calendar or fiscal year having due regard to the reasonable profits which, but for such arrangement or understanding, might or could have been obtained from dealing in such products, goods or commodities."[18]

It is submitted that this provision is too restrictive in that it makes necessary a showing of "profit-diverting devices" before the Department can require consolidated statements: the Department should have the power to require consolidated statements in all cases of unitary business performed by two or more commonly controlled legal entities; and the allocation of income attributable to Wisconsin should be based upon such consolidated statements.

Apportionment Factors

As previously noted, Wisconsin employs three factors in its allocation formula: tangible property owned and used; cost of manufacturing; and sales at the office by or through which sales are accepted.   The discussion following provides a critical evaluation of these factors as defined in the Wisconsin statutes together with some suggestions for revision.

---

[14] The question of "business situs" is not considered here. It may be that as the "business situs" of intangibles is more clearly defined some exception to domicile allocation of intangible non-apportionable income should be made. Patent or copyright royalties might be allocated to the state where the rights are utilized, unless this can not be established, in which case they would be allocated to the state of commercial domicile.

[15] Section 71.07 (2) Wisconsin Statutes 1957.

[16] For a discussion of unitary business see, John A. Wilkie, "Unitary Allocation of Income from Unitary Business," *Taxes*, May 1959.

[17] See an analysis of separate accounting by George I. Altman and Frank Keesling, *Allocation of Income in State Taxation*, 2nd edition (Chicago, Illinois Commerce Clearing House Inc., 1950), pp. 88-97.

[18] Section 71.11 (7) (a) Wisconsin Statutes 1957.

Property Factor: Inclusion of Rental Property. Wisconsin law provides that the tangible property factor used in the apportionment formula applies only to that tangible property real, personal, or mixed, "which is *owned and used* by the taxpayer in Wisconsin in connection with his trade or business during the income year" and "to such property of the taxpayer *owned and used* by him in connection with his trade or business everywhere." It is submitted that rental property which is used by the business should be included—as is now the case in several states.[19]

Exclusion of rental property from the property factor can result in discrimination against firms that own all property which they use in Wisconsin. For example, assume that X and Y corporations do exactly the same amount of business and that each has cost of manufacturing and sales factors of 50 percent. If X owns its buildings in Wisconsin and neither owns nor rents property elsewhere its property factor would be 100 percent. On the other hand, if Y rents its buildings in Wisconsin and owns no property in the state its property factor would be zero. Thus, X would apportion two-thirds of its income to the state of Wisconsin while Y, which does the same amount of business in Wisconsin, would apportion only one-third.[20] Similarly, a corporation with rented property outside the taxing state could pay more taxes than a similar corporation which owned its property outside the state. Valuation of rented property could be on the basis of

eight times the net annual rental. This ratio is used by most of the states which include rental property.

Cost Factor: Substitution of Payroll for Manufacturing Cost. Wisconsin law provides for a "cost of manufacturing, collecting, assembling, or processing" factor in the apportionment formula.[21] In the opinion of the author this factor should be replaced by a payroll factor, and for a number of reasons. First, it should be noted that payroll cost is present in all phases of the business. It is present not only in manufacturing but in purchasing, research and development, designing, storing, transporting, advertising, selling, general administration, policy determination, etc. The effect of combining a payroll rather than a cost-of-manufacturing factor with the property and sales factors is to give less weight in the formula to plant locations where the manufacturing expenditures are concentrated and more weight to the location of overhead and/or marketing operations.[22]

Secondly, there is less overlap between factors where payroll rather than manufacturing cost is used. Thus the cost of depreciable assets does not appear in a payroll factor but is part of both manufacturing cost and property. Some of the costs that are part of the manufacturing cost factor (e.g., profits attributable to an increase in semi-processed and finished goods inventories) would also overlap with components of the sales factor. This would not be the case with a payroll factor.

Thirdly, the major inequity of the payroll factor is in certain cases also the major inequity of the cost of manufacturing factor. The differential of real wages between operations in different states may not be reflected in differences in

---

[19] This is recommended by the National Conference of Commissioners on Uniform State Laws of the American Bar Association, Sec. 10, Uniform Division of Income for Tax Purposes Act, *Taxes*, August 1957, p. 632, and in the Final Report of the Committee on Tax Situs and Allocation *1951 Proceedings National Tax Association*, Dallas, Texas, p. 461. Among the states that include rental property are New York, Delaware, Maryland, Montana, Tennessee, Oregon, Vermont and Minnestoa.

[20] Adapted from "Uniform Division of Income for State Tax Purposes," William J. Pierce, *Taxes* October 1957.

[21] Section 71.07 (2) (b) Wisconsin Statutes 1957.

[22] No estimate was made of the fiscal effect of adopting a payroll factor in lieu of the cost of manufacturing factor.

productivity. Higher wage rates may actually mean higher unit cost of the product produced for sale in the plant in one state compared to that in another.

Fourthly, the payroll factor is not limited to unitary manufacturing concerns. It is quite appropriate for merchandising concerns, whether wholesale or retail, personal service businesses and even certain public utility businesses. The manufacturing cost factor, however, is confined to those concerns engaged in the manufacture or production of goods.

Fifthly, the payroll factor applies to "resale" activities of manufacturing concerns. Manufacturing costs cover only manufacturing operations and thus should be weighted where the unitary business involves any "resale" operations whether at wholesale or retail level. Such weighting requires additional compliance and administrative costs which is not necessary for the payroll factor. When joint costs are involved there is the problem of allocating such costs between manufacturing and selling of the firm's own products and its "resale" operations.

Sixthly, the payroll factor is much less difficult to administer than the "manufacturing costs" factor. The application of the cost of manufacturing calls for allocation not only of overhead but of costs of raw materials. This may necessitate an arbitrary allocation of joint costs of raw materials between different production plant operations in different states, and reliance on average instead of actual cost of goods and supplies may be used in the manufacturing factor. There is a more simple assignment of overhead payroll than of "manufacturing burden" or "overhead" to various manufacturing, assembling, processing and collecting activities within a state. The cost of manufacturing factor calls for numerous audit adjustments of certain items which are not involved in the pay-

roll factor such as federal excise taxes, depreciation, and shipping expenses. The payroll factor can be computed rather easily by reference to unemployment insurance data.[23]

Seventhly, the payroll factor is more widely used as a part of a three-factor formula for manufacturing concerns than is the cost of manufacturing factor. In recent years the number of states using the payroll factor has increased whereas the number using the manufacturing cost factor has decreased. The trend is definitely toward the three factor formula which includes the payroll factor.

Sales Factor: Adoption of a Destination Basis. Wisconsin law provides as a third apportionment factor a sales ratio based upon "total sales made" through or by offices, agencies, or branches located in Wisconsin during the income year as related to the total net sales made everywhere during such income year.[24] It is submitted that this basis should be changed to a destination basis as suggested by the recent National Conference of Commissioners on Uniform Laws.[25] There it was "believed that the contribution of the consumer states toward the production of the income should be recognized by attributing the sales to those states."[26] We are concerned here with taxing net income of the taxpayer and not with the taxation of activity as such. In an exchange economy, income of the seller is dependent upon the effective demand of the buyer. Not only must there be a desire for the product or service but there must also be the money

[23] This data may have to be adjusted with respect to some employees who actually perform services in more than one state. Equitably speaking, such salaries and wages should be allocated upon a time basis. If required this may add some compliance work to such a factor.

[24] See 71.07 (2) (c) Wisconsin Statutes 1957.

[25] See *Taxes* August 1957. It was also recommended in the Final Report of the Committee on Tax Situs and Allocation 1951, *Proceedings National Tax Association*, Dallas, Texas, p. 463.

[26] William J. Pierce, "The Uniform Division of Income for State Tax Purposes," *Taxes*, October 1957.

or its equivalent available to provide payment for such goods. Effective demand may be localized at the place where the product or services are received by the customer which in most cases will also be where they are consumed. It should be recognized, however, that effective demand for a given corporation's products or services may be considerably influenced by the corporation's own activities both in production and in distribution (marketing efforts). The effective demand of the consumer, as localized geographically, is dependent upon certain services that are provided by the state government in such areas. Without such benefits from government the effective demand would not have been developed and been capable of current realization. The effect of giving no weight to the place of market consumption is to give a tax advantage to the multistate business over the local business in an income tax state, or it may be giving tax advantage to one unitary corporation over another.

The receipts factor on a destination basis acts as a balance to the property and payroll factors. It does so even better than the other methods based upon "selling activity"--which are also represented in the property and payroll factors. A sales factor based upon the office of acceptance or rejection of the sale gives additional weight to the home office state where the plant (or main plant) is located.

A receipts factor based upon destination also provides little opportunity to affect tax liability by locating capital and labor relating to marketing operations in one state rather than another. Thus, such installations and offices whether owned or rented would be located geographically more on the basis of economic factors without considering special tax effects. Where other factors are approximately equal these special tax effects can be of some importance in the decisions on location of operations especially where the offices, etc., are of a character that they can be easily moved from one place to another (such as "sales offices").

The main argument against the destination basis is that under current concepts of what constitutes "doing business" within a state for income tax purposes some receipts from sales may be apportioned to states which have no tested power to tax any net income which is partly derived from such a source.[27] In such cases, however, the sales involved could be attributed to another state upon the basis of some other standard such as the state from which the goods are shipped,[28] or the state where the sales activity is performed, or the state where the sales are "chiefly negotiated," or to the state where the sales are accepted or rejected. Which of these is preferable is not readily determined. The state from which goods are shipped rule, while clear in most cases, provides that where the goods are not located at a permanent place of business of the taxpayer outside the state but somewhere else the sales should be assigned by reference to an "office situs" rule. This provision again is designed to prevent sales from being assigned to a place where the taxpayer is not doing business. It may be more simple to provide for an "office situs" rule to cover all such cases. The point of shipment is already represented in the formula in the property factor. The receipts factor might well represent the location of most of the "selling activity"

---

[27] The author submits that such power may be constitutionally supported since a benefit is offered to the taxpayer at the place of consumption.

[28] This is recommended by the National Conference of Commissioners on Uniform Laws. See, *Taxes*, August 1957. It is also recommended in the "Final Report of the Committee on Tax Situs and Allocation," 1951, *Proceedings National Tax Association*, Dallas, Texas, pp. 463-4.

of the corporation. Such "selling activity" may not be sufficient within a state to warrant tax jurisdiction under current concepts of doing business. Thus, exceptional sales would again have to be allocated on the basis of some office situs rule. Although it may be somewhat arbitrary in some cases it is submitted that, because of the greater degree of certainty, the office, agency, or branch of acceptance or rejection be used.[29] Like other methods this may be subject to some manipulation to avoid taxation. Such attempts when clear should not be countenanced.

The destination basis may also be objected to as to its treatment of exports. Should such sales be allocated abroad? There are two types of exports, those which are commercial and those in which the purchaser is the United States Government. Commercial exports could be treated similar to domestic sales. When such sales are allocated to a country which has no basis for taxing the corporation (under current doctrine) they could be allocated to the state where the office of acceptance or rejection is located. In the case of sales to the United States Government for delivery to points of embarkation such sales could also be allocated on the basis of the state where the sale was accepted or rejected.[30]

As for the fiscal effects of the destination basis it may be observed that the predominantly "manufacturing" states are also large consumption states and that the adoption of the destination rule would not effect revenue substantially.[31]

[29] The office where the sale is "chiefly negotiated" is not recommended because of the uncertainty of the identifying element in determining the office where a sale is "chiefly negotiated and executed" or the office, etc., from which the salesmen mainly responsible function.

[30] The National Conference of Commissioners on Uniform Laws recommends the state from which the goods are shipped. See, *Taxes*, August 1957.

[31] It might be advisable to make an estimate of the fiscal effect particularly among the large taxpayer corporations.

It might be argued that some other basis than "destination" should be considered in formulating the sales factor. One might be that of "sales activity" such as that used in California, namely, the location of employee sales activity or where orders are solicited and obtained by salesmen. This method in effect approaches closer to the "destination" basis than any office situs method or point of shipment method since the state of solicitation and promotion is apt to be the same as the state of destination of the product. This factor is not a very good balancing factor since the services of employees engaged in selling activities are already represented in the payroll factor. Receipts from sales which cannot be allocated to states without jurisdiction to tax would have to be allocated on some other basis such as the office situs rule. There are also many administrative problems where the sales activity in connection with a sale is performed in more than one state. What standard can be formulated for uniform apportionment of such sales? It might be on the basis of where the most time was spent or where the greatest value of the services is performed.

Another method for allocating sales is on the basis of the sales office or other place of business at which the sales are chiefly negotiated or from which the salesmen mainly responsible for the sale operate. In favor of this method it is argued: (1) that such office is representative of the location of the "selling activities" of the corporation and should be treated as a factor of major importance in the production of apportionable income; (2) that the method acts as a balance to the property and payroll factors; (3) that all sales are allocated to jurisdictions able to tax; and (4) that there is little if any effect of such method upon the sales office location. However,

this method has these weaknesses: (1) it does not recognize the contribution to net income of market consumption or the place of effective demand; (2) "selling activities" of a unitary manufacturing or merchandising corporation should not be treated separately as a major source of net apportionable income (they are already represented in the property and payroll factors); (3) "selling activities" if used should not be limited merely to those connected with obtaining sales or transactions; (4) the office from which the salesmen function may not be representative of the selling activity of the corporation; (5) there is considerable uncertainty of the identifying element in determining the office, etc., where the sale is "chiefly negotiated and executed," or the office from which the salesman mainly responsible functions; and (6) in certain situations such as a nearby city in a non-tax state the identifying element is too easily shifted outside the state. On balance, it is concluded that this method is not preferable to the "destination" basis.

Another major type of office situs allocation rule for receipts from sales is based upon some particular act in connection with the sales, such as the office where the order is accepted. If the sale is allocated at the office where the sale is accepted it should not be allocated to any other jurisdiction on the basis of some office through which the order was obtained or received. The strong argument for this type of method is that the place of acceptance or approval of a transaction is usually readily identifiable. It not only has judicial sanction but gives recognition to the location of the head office in most cases and gives effect to the sales effort and policy of the head office. However, this method (1) gives no recognition to the importance of market consumption; (2) in many instances pro-

vides no balancing effect to the property and payroll factors; (3) does not recognize the many selling activities located away from the accepting office; (4) works unneutral effects upon similar competitors purely on the basis of where the approving office is located; (5) has some administrative difficulties in determining where the order is accepted or rejected and how to handle the order where elements of acceptance are handled in separate offices in separate jurisdictions; (6) is readily avoided by shifting the accepting office to a non-taxing or lower taxing state; and (7) is based upon legal decisions which by and large are based upon the idea that the sale itself is being taxed rather than the net income derived partly therefrom. Here again it is submitted by the author that this method is not preferable to the "destination" basis of allocating sales for apportionment purposes.

The last alternative considered is the method of allocating sales on the basis of the place from which the goods are shipped.[32] This method is said to give representation to the "business "activity of the firm. A warehouse or place of business from which goods are shipped is located in a given state because the location facilitates the earning of net income or has a distinct bearing upon the earning of such net income. The method, however, if adopted would ignore the role of market-consumption or effective demand. The states of origin of shipment differ in many cases from the states of location of the customer. This method does not act as a balancing factor to the property and sales factor. It is to some extent represented in the property factor. Nor is the place of shipment necessarily the location of "selling activity" of major importance to the sale. As stated above,

[32] This method is used in Alabama, New York, and Oregon.

the rule would have to be compromised where the goods are not located at a permanent place of business of the taxpayer but somewhere else and would probably be assigned by reference to an "office situs" rule. This would be done to prevent a "no man's land" of sales being allocated to states without jurisdiction to tax. This might be remedied by the expansion of the doing-business concept to include the location of any property whatsoever in a state as being sufficient basis for jurisdiction to tax.[33] This method, it is concluded by the author, is not preferable to the "destination" basis of allocating sales for apportionment purposes.

On the basis of the above analysis of various approaches to allocating sales, it is concluded that the preferable method is that based upon destination. The complete adoption of this method may not be feasible immediately because of limitations currently being put upon the concept of "doing business' (as noticed above by the modification with an office situs rule). Nevertheless, it should remain the objective to be striven for and the direction for concentrating efforts to work out the administrative and enforcement problems related thereto. At the same time the concept of "doing business" should be expanded to include the more direct and indirect activities of the taxpayer or its agent or representative. Receipts from services rendered in connection with unitary business operations should be allocated for apportionment purposes on the basis of where such services are performed.[34]

Rather than use of a sales factor other than one based on destination, it is suggested that the sales factor be omitted altogether.[35] This is based on the argument that the activity of the corporation including the sales activities is already adequately represented in the property and payroll factors.[36]

Conclusion

In conclusion, it is suggested that Wisconsin's allocation procedure could be substantially improved with the adoption of the following changes in the law: (1) broaden the basis of tax jurisdiction;

---

activity is performed in more than one state, other sales are attributed to the state in which the greater proportion of the activity was performed, based upon costs of performance. In many types of service functions this approach appears adequate. However, there are many unusual fact situations connected with this type of income and probably the general provisions of Sec. 18 should be utilized for these cases. If we assume the activity involved is the servicing of industrial equipment, the formula provided in the uniform act could be easily applied and the result appears equitable. In contrast assume that the sales action involved is advertising revenue received by a national magazine publisher. The state of activity would be difficult if not impossible to ascertain so it would appear that this type of income may well be apportioned on the same basis as subscription income. The national conference considered this problem at length and concluded that, for certain types of sales income, exceptions would have to be established by the tax collection agencies since no formula seemed to be satisfactory for every conceivable factual situation." See, William J. Pierce, "The Uniform Division of Income for State Tax Purposes," *Taxes*, October 1957.

[35] A Wisconsin Tax Commissioner has stated: "As I see it the sales factor is based upon a false premise. It presumes that the income produced by the sale has its origin in one state where as a matter of fact, the efforts of people in three or four different states probably produced the sale. A traveling salesman, might have picked up the order, sent it through a regional office in a second state and on for final approval in a third state. But the sales factor presupposes that the sale earned income in only one of these three states. It is a false premise and no matter how you define the sales factor you are short-changing one of those three states . . . . It has always seemed to me that a uniform formula based upon property and payrolls would allocate income to the state in which that income is earned." Harry W. Harder, 1956 *Proceedings National Tax Association*, Los Angeles, California, p. 173.

[36] Another proposal is to use a uniform tripartite sales factor which would include negotiation, origin, and destination with equal weight for each factor. This is recommended as being most equitable and probably bringing the smallest overall change in state revenues. It would certainly "complicate the accounting work for the corporations and the auditing work of tax administrators." Paul M. Holt, "Interstate Allocation of Corporate Income," 1956 *Proceedings, National Tax Association*, Los Angeles, California, p. 168.

---

[33] Such as goods on consignment with brokers or independent dealers, or goods in a public warehouse, or even samples used by salemen or their agents.

[34] This is recommended in the "Final Report of the Committee on Tax Situs and Allocation," 1951 *Proceedings National Tax Association*, p. 463 and by the National Conference of Commissioners on Uniform State Laws, *Taxes*, August 1957. In explaining the provision adopted by the National Conference Professor William J. Pierce declares: "If the

(2) adopt commercial domicile of the recipient as the basis of allocating intangible non-apportionable income of corporations; (3) eliminate separate accounting for unitary businesses; (4) provide for a consolidated statement basis of apportionment for all unitary businesses of two or more commonly controlled legal entities; (5) include rental property in the property apportionment factor; (6) adopt a payroll apportionment factor in lieu of the manufacturing cost factor; and (7) adopt a sales factor based primarily on destination.

On the Output Unit in Transportation

By GEORGE W. WILSON*

THE PROBLEM of a proper conception of output units is one which cuts across large segments of economic analysis. In the labor market, certain product markets and service industries, the fact of non-homogeneity creates serious problems of analysis which are frequently though happily not always assumed away. However, in the field of transportation economics there has been a general disposition to neglect the nature of the transportation product and to presuppose homogeneity at least for the product of particular media of transport. This presents more than analytical difficulties. It has important practical consequences for, if it is true that an optimum transport system requires rates to reflect costs, then a clear conception of the product or thing to be "costed" is an essential precondition. Until we are clear on the nature of the output of transportation firms we are scarcely in a position to assign costs and thus optimum allocation may be thwarted. My purpose here is to analyze some of the problems involved in assessing the unit of transportation output.

The question of the homogeneity of the unit of transportation output seems to have been largely neglected since the Pigou-Taussig controversy.[1] Recently, however, the issue has been raised once again by several authors. Barger asks: "Is the ton, the carload, the ton-mile, the carload mile, a combination of these, or some altogether different measure to be taken as the fundamental unit of service rendered in the realm of freight transportation?"[2] Unfortunately, he fails to get beyond the question and arbitrarily adopts the weighted ton-mile.

Two more recent studies have emphasized the heterogeneity of the ton-mile and have suggested alternative conceptions. Milne argues: "It is highly misleading to regard all transport facilities as parts of one industry, the transport industry, and as producing homogeneous passenger miles . . . . and homogeneous ton-miles. The characteristics of the demand for transport services is its heterogeneity.[3] Later, Milne suggests: "It is more convenient to deal in terms of what we have called transport units and to regard the train-journey, the bus-journey, the truck-journey . . . as our unit of output." (p. 121-22) Milne would furthermore distinguish this from "the pricing unit" which he construes as the "individual passenger and the individual consignment." (p. 125) Thus, the unit in which to reckon costs is quite different from the unit in which to reckon sales or demand.

In a similar vein, Troxel, after raising the question, "What is the definition of the output or commodity unit?"[4] suggests that some of the trouble about output units follows:

". . . . from a habitual attention to the output units in which sales or exchanges are made . . . [The] pricing diagrams of economists commonly are the same quantity units . . . . for both cost and demand relations. Showing similarly an attention to 'final' or

* Associate Professor of Transportation, Indiana University, Bloomington, Indiana.

[1] A. C. Pigou, *The Economics of Welfare*, 4th ed. (London, England: Macmillan and Company, Limited, 1950), Chapters XVII and XVIII; F. W. Taussig, "Railway Rates and Joint Cost Once More," *Quarterly Journal of Economics*, May 1913, p. 378; and F. W. Taussig with A. C. Pigou, "Railway Rates and Joint Cost," *ibid.*, August 1913, pp. 535 and 687.

[2] Harold Barger, *The Transportation Industries, 1889-1946* (New York: National Bureau of Economic Research, 1951); p. 176.

[3] A. M. Milne, *The Economics of Inland Transport* (London, England: Pitman, 1955), p. 83.

[4] Emery Troxel, *Economics of Transport* (New York: Rinehart & Company, 1955), p. 93.

exchangeable outputs . . . . students of transportation refer habitually to such units as ton-miles, passenger-miles, carloads and truckloads. Only a small part of total costs are variable directly . . . . in relation to these 'final' products. Indeed, the organization of transport operations is not much embraced in ton or ton-mile, passenger-mile, or even load units." [5]

The implication of Troxel's argument is that ton-miles may be homogeneous on the demand side but not on the cost side whereas, if I read Milne correctly, ton-miles are not homogeneous on either side. There is clearly a good deal of confusion on this point which merits some attempt at clarification.

Let us begin by reviewing briefly the pertinent aspects of the discussion between Pigou and Taussig in 1913. [6] The argument centered on the question of whether railway pricing was monopolistic in origin or simply a manifestation of joint supply with Pigou upholding the former and Taussig the latter proposition. In Taussig's words: " . . . . railways always supply 'transport' . . . . In this sense they supply a single sort of thing, or a homogeneous commodity. But I submit that they do not supply commodities or services which are homogeneous in the sense important for the purpose in hand,—namely, as regards the conditions of demand." [7]

In short, if the demand for transport differs among the various shippers, then we have the "peculiarities resulting from joint supply." [8] Taussig used the production of cotton fiber and cotton as analogous to the movement of two commodities X and Y. Obviously the former is a case of "true jointness" de-

fined as a situation in which "the production of one commodity *necessarily* results in the production of another." [9] But it is not so readily apparent that the service of transporting one commodity *necessarily* requires the transportation of another. Pigou holds that these services are not joint and that, under conditions of competition, identical rates would be charged for their movement. [10] The question then became diverted from the nature of the unit of transport output to the definition of joint cost. In this latter aspect there was no doubt that Pigou had the most serviceable definition and drew clearly and correctly the distinction between joint and common costs. [11] As a result the consensus supported Pigou, although Allyn Young was charitable enough to see: "no inconsistency in granting both Professor Taussig's contention that railway costs are largely joint costs and Professor Pigou's contention that railway rates form a special case of discriminating monopoly price." [12] Yet surely the issue encompassed more than the matter of joint costs. For if transportation firms were multiproduct enterprises (i.e., if the output were not homogeneous ton-miles or homogeneous "transport") then competition would hardly equate the prices of different products. Competition would, of course, equate prices to the marginal costs of the various outputs and thus there would be no discrimination although prices would differ. But this is not the situation envisaged by Pigou. Pigou held to the view that the unit of transport service as

[5] *Ibid.*, pp. 93-94. A. E. T. Griffiths has labelled the assumption of homogeneity as "quite clearly ridiculous," in "Cost Variations in Transport," *British Transport Review*, August 1957, p. 470.

[6] See footnote 1.

[7] "Railway Rates and Joint Cost Once More," *Quarterly Journal of Economics*, May 1913, p. 381.

[8] *Ibid.*

[9] D. P. Locklin, "The Literature of Railway Rate Theory," *Quarterly Journal of Economics*, February 1933, p. 194. Viner has construed joint costs in a similar fashion in *Encyclopedia of the Social Sciences*, Vol. II (New York: The Macmillan Company, 1937), p. 473.

[10] "Railway Rates and Joint Cost," *Quarterly Journal of Economics*, August 1913, p. 689.

[11] See also J. M. Clark, *Economics of Overhead Costs* (Chicago, Illinois: University of Chicago Press, 1923), pp. 58-59.

[12] "Pigou's Wealth and Welfare," *Quarterly Journal of Economics*, August 1913, p. 681., fn. 2.

applied to different commodities was homogeneous on the supply (i.e. cost) side although for the immediate discussion he (temporarily) accepted Taussig's presumption of nonhomogeneity thereby turning the whole issue into one of joint costs (and its definition) versus monopoly power as an explanation of railway rates. But Pigou's victory on the score of cost definition and monopoly power apparently carried with it acceptance of homogeneity of transport output on the supply side. Yet surely the two are separable. Nonetheless, the question of the nature of the unit of output was apparently resolved in favor of homogeneity until most recently. Winning on only one point, Pigou took all the chips. There are, then, two main problems regarding the unit of output: (1) What is it? and (2) Is it homogeneous?

Transportation firms create "place utility." That is, they move things from points where their economic value is less to places where their economic value is greater. Transport firms therefore produce a product which is bound up with weight and distance. There can be no question then that ton-miles appropriately encompass the function of transport. But what of Troxel's and Milne's contention that the ton-mile is a *sales* unit which differs from the cost-unit? Clearly the usual conception does emphasize the sales side. Is this legitimate?

An optimum transport system requires that rates be based upon costs. Rates are, of course, the prices paid for the "final" product and must be construed in terms of sales units. If rates are to reflect costs it becomes clear that, whether transport equipment is geared to the ton-mile or not, a costing technique running in terms of the final product is essential. Otherwise, we revert to value of service pricing which, especially where inter-carrier competition is concerned, is

not only untenable but with freedom of shipper choice leads to resource misallocation.[13] That is, suppose we make up a train to run between A and B and suppose we adopt Milne's supply unit (i.e., the "train-journey"). We could then determine the cost of a train-journey, A-B, say $K. But we transport a variety of goods, l, m, n, ... z, and must therefore establish a price for moving each. In the long run the total cost of running the train A-B and return must be recovered. If we cannot assign *costs* to each of l, m, n, . . . z, we must engage in value of service pricing and hope that the aggregate receipts therefrom exceed long-run total costs. Note, however, that if this is the only reasonable way in which to assign costs (i.e., in terms of units which are much in excess of the pricing unit) then we can never get a cost based rate *structure*. We can, of course, obtain a cost based rate level but no more. If the views of Troxel and Milne are correct, all the recent emphasis upon equating rates with costs has been missing the essence of the economics of transportation. But if we can assign costs on any reasonable basis to the sales units then we must do so else the national transportation policy cannot be fulfilled. The point here is that even though the supply unit may differ from the sales unit it is clear that there is no choice but to adopt the latter and ascertain costs in these terms. This is especially true if we insist on a cost-based rate structure since rates refer to sales units. No railroad offers a train-journey for sale. There is therefore little point in devising accounting techniques in these terms unless any other basis is completely unwarranted. Thus we are compelled to adopt the sales unit as relevant to *both* the supply (cost) and demand sides. In terms of the transport

[13] See J. R. Sargent, *British Transport Policy* (Oxford, England: Oxford University Press, 1958).

function, as argued above, the ton-mile is the most appropriate unit especially with reference to all forms of transport.

But having decided upon the ton-mile we must face up to the issue of homogeneity. There are two aspects which must be analyzed: (1) homogeneity on the demand side and (2) homogeneity on the supply side. The ton-mile is a product of two variables, weight and distance, each of which, taken separately, is homogeneous. But weight is one of the physical properties of a commodity. When the commodities differ does this make ton-miles heterogeneous? That is, since a ton of oranges differs in an important physical and economic sense from a ton of coal, then does not a ton-mile of service applied to oranges differ from a ton-mile of service applied to coal? On this point Locklin has joined with Taussig holding that "lack of homogeneity may not be due to physical differences in two commodities but to a difference in demand for them. Usually the two go together but not always."[14] If Taussig and Locklin are correct then the ton-mile is a highly variable product changing with every change in the item to which it is applied. Under this conception each transportation agency would be construed as a multiproduct firm producing ton-miles of l, m, n, . . . z, etc., which would in turn make cost finding extremely complex.

Now there is no argument that in a "physical" sense ton-miles are homogeneous or in Taussig's words, that carriers supply a form of homogeneous "transport." (Isard calls these "transportation inputs.")[15] The question therefore reduces to the following: Do different demands for physically identical units of output render the units economically heterogeneous? To answer this it should be pointed out that the transportation of different commodities represents simply different uses of transport facilities or, more accurately, sales of ton-miles of service to different shippers. The level and slope of individual demand schedules for any particular product has never *per se* sufficed to alter the nature of the product. Otherwise, there could be no such thing, even a priori, as a perfectly competitive market unless one argued that the level and elasticity of each buyer's demand function were identical. However, while different individual uses of (or demands for) a particular product do not change the nature of the product, they create an essential (though not sufficient) condition for price discrimination. But even where different prices are charged this does not alter the case. For example, when one ton of X and Y are each moved 100 miles this is merely the sale of 100 ton miles of service to different buyers (the shippers of X and Y) analogous to the sale of *n* tons of a certain grade of steel to an automobile manufacturer or a construction firm. If the automobile manufacturer receives a better price than the construction firm, no one would argue that the steel was any different. Why then hold that ton-miles of X is distinct from ton-miles of Y solely on the grounds that X and Y are different? As Machlup puts it, ". . . . discrimination in railroad transportation . . . . refers to one product, namely transportation service, which the railroad offers at discriminatory prices to different groups, namely persons using the service for different commodities."[16]

But while this takes care of the allegation that different demands per se can create heterogeneity it does not com-

[14] D. P. Locklin, "The Literature of Railway Rate Theory," *Quarterly Journal of Economics*, February 1933, p. 189.

[15] *Location and Space-Economy* (Cambridge, Massachusetts: Technology Press of Massachusetts Institute of Technology, 1956).

[16] Fritz Machlup, *The Political Economy of Monopoly* (Baltimore, Maryland: Johns Hopkins Press, 1952), p. 136.

pletely resolve the problem. For example, if it took five hours to move one ton of X 100 miles but two days to move one ton of Y 100 miles, would we then be justified in calling ton-miles of X and ton-miles of Y homogeneous? This would at least create a difference in quality of service as would also considerations of door-to-door delivery (i.e., flexibility of service), dependability and safety. This is especially important as regards ton-miles generated by different media or by the same medium in different localities and under different conditions. In short, such product differentiation raises the whole question of how to define a product or industry. How much of such differentiation can there be before similar products become dissimilar enough to warrant treating them as distinct and produced by firms in a different industry? Is there such a thing as *the* transportation industry?

Again we have widely divergent views. Isard refers to "the transport rate" which equilibrates the supply of and demand for transportation inputs.[17] This broad conception (also held by Pigou and Machlup) may be contrasted with that of Milne and Pegrum. Pegrum has stated:

"Land transport cannot be regarded as a homogeneous industry in any respect. It cannot even be regarded as an industry. Instead it is a group of industries composed of agencies such as railroads . . . pipelines and automatic vehicles . . . These different agencies have diverse economic characteristics. They supply many different kinds of services, some competitive, some complementary and some quite different. They offer these services under markedly different technical conditions and very different market conditions."[18]

[17] *Op. cit.*, p. 86, Walter Isard has elsewhere suggested that "one can utilize the Marshallian approach (in transportation) if he follows a *functional* analysis." (See, "Distance Inputs and the Space-Economy, Part 1: The Conceptual Framework," *Quarterly Journal of Economics*, November 1951, p. 196. In *Location and Space Economy*, Isard substitutes "transportation inputs" for "distance inputs."

[18] Dudley Pegrum, "The Economic Basis of Public Policy for Motor Transport," *Land Economics*, August 1952, p. 245,

How can these views be reconciled? Are we to accept the more generalized or the more specialized view of the industry, both of which depend upon whether one construes the unit of output as homogeneous or not as among the various media? (Note that heterogeneity here does *not* refer to different commodities, only to different media.)

For purposes of general investigation, such as trying to develop theories of location or integrate a distance factor into the theory of the firm, the generalized conception is appropriate. Any other view would overly complicate the basic purpose. Thus, as is so often the case, we utilize various simplifying though in this case unspecified assumptions and ignore some of the less essential complications in order to get on with the major task. On the other hand, as soon as the major purpose relates directly to the provision of transport the complications must be considered. It is not completely fortuitous that Milne and Pegrum are known as "transportation" economists whereas Pigou and Machlup are known best for their broader theoretical work. In short, whichever conception of transportation is deemed appropriate depends upon the purpose; and, generally speaking, the broader the analytical vista the more appropriate is the more generalized view of the transportation industry and vice versa.

But the whole problem of whether transportation is one industry or many involves the basic question of relative substitutability (and hence homogeneity) among the services of the several media which in turn requires analysis of the cross elasticity of demand as well as available alternatives. In this respect, demand has an important role to play though in a different sense than Taussig

and *Price Competition in Transportation* (Chicago, Illinois: Railway Progress Institute, 1956), p. 15.

or Locklin had in mind. Likewise, if the above generalization is accurate, then any discussion of inherent advantages must consider inter-carrier differences (and similarities) and hence cannot avoid tackling the problem of relative substitutability (i.e., one industry or many?).

We define an industry as one in which the cross elasticity of demand for the "product" vis a vis other "products" is low[19] so that if *all* the firms in this industry raised (or lowered) their prices by the same amount (ceteris paribus) the substitution effect (as distinguished from the income effect) would be negligible. That is to say, after abstracting from the income effect, the elasticity of demand for the product would be low. On the other hand, the cross elasticity of demand for the product of those firms comprising the industry is very high (again omitting any income effects). In the limiting case where the goods are homogeneous or non-differentiated, the cross elasticity is infinite—the case of perfect competition. Using this to test the Milne-Pegrum thesis, would a rate decrease, say, by all motor carriers for a particular product (ignore private trucking) lead to a negligible or substantial increase in demand for trucking services? The answer depends upon the commodity involved, whether the short or the long run is the relevant time span, and specific shipper circumstances.

That is, for some shippers, depending upon commodity, location or specific circumstances, there will be distinct savings through utilization of one medium over another in addition to any freight rate differences. The "costs" to the shipper of using a particular form of transport therefore depends upon the freight rate actually charged by the transport firm and the *quality* of the transport service. By the term quality of service we refer to those elements in the provision of ton-miles which amount to a form of product differentiation such as speed, flexibility, safety and dependability. It is obvious that improvements in each of these qualitative aspects will generate non-transport cost savings of varying amounts for particular shippers. More rapid transit (if coupled with dependability) permits inventory reduction and hence a saving in storage costs. Greater flexibility, in the sense of door-to-door delivery, not only reduces total transit time but means less handling and thus greater safety and dependability in the sense of increased assurance of prompt arrival of the consignment in good condition.

It is apparent that the magnitude of any non-transportation cost economies to be derived from varying service qualities will vary substantially among shippers, depending upon location, type of commodity and inventory policy (the latter is liable to periodic fluctuation depending upon the general condition and prospects for the industry). For example, if a shipper is located along a rail line, the quality of rail service will be greater to him than to some other shipper not having direct access to rail service. Again, shippers of different commodities will attach varying degrees of importance to quality elements. Goods which are not susceptible to damage in transit and/or which can be stored easily (e.g., many raw materials) will not value very highly the qualitative elements. On the other hand more fragile or perishable items, especially those whose market price fluctuates, will value quality of service very greatly.[20]

What this all means for transportation is simply that the degree of elasticity of

[19] See J. S. Bain, *Pricing, Distribution and Employment*, Rev. ed. (New York: Henry Holt & Company, 1953), pp. 23-26 and 50-52.

[20] For a fuller discussion of these points see Milne, *op. cit.*, Chapter IV.

demand for the service of any particular medium of transport depends upon the "net cost" of using that particular medium. By net cost is meant the freight rate minus non-transport cost savings due to the value which any shipper places upon the quality of service rendered by any particular transport medium. But this latter valuation is not only highly relative (depending upon circumstance) but also highly variable over time. Thus the degree of substitutability among forms of transport is likewise extremely variable and volatile depending upon particular commodities and shipper circumstance at different times. Under our definition of an industry, therefore, it is apparent that each of the several forms of transport cannot constitute a separate industry. The degree of substitutability cuts across the different forms and does not depend uniquely upon any single one.

Identification of an industry by the particular form of transport is too narrow. Substitute conditions are no respecters of technical conditions of supply and cut across the various media on the basis of price-quality considerations which in turn depend upon commodity, location and circumstance. On the other hand, construing transportation as a single industry is too broad for purposes of analyzing inherent advantages within transportation. In terms of the unit of output, this means that the product (ton-miles) is "honogeneous" only insofar as substitutability is high. We can do little better than this on the demand side especially since the range of substitutability varies with changing conditions. Thus, on the demand side we can suggest that different qualities of the service, where they are such as to create a substantial range of demand inelasticity, likewise induce heterogeneity.

Where the qualities are not of this nature, ton-miles are more nearly homogeneous. In this sense Taussig and Milne are correct—demand conditions can create heterogeneity—but there is no inconsistency in granting Pigou's contention either.

On the cost side it must be emphasized that transportation output is a product of at least three variables—weight, distance and velocity. If costs behaved in the same fashion with respect to changes in each of these three variables there would be no problem and the unit of output would be homogeneous. But the fact is that each of the variables is an independent source of cost and thus different combinations of them involve different costs. For example, ignoring velocity, 100 ton-miles may represent 1 ton moved 100 miles, 4 tons moved 25 miles, 100 tons moved 1 mile (etc.). Since each of these represents different combinations of "loading" and "moving" inputs, it is evident that the costs for 100 ton-miles must differ, depending on the proportion of tons and miles involved. We need therefore to consider cost variability with respect to varying combinations of weight, distance and velocity. McDonald has expressed the matter thus:

"Only when two out of three qualitative elements are given can we distinguish a homogeneous product . . . Thus if we plot weight, distance, and rate of output on a three-dimensional diagram total cost cannot be shown but will vary uniquely as we move parallel to one axis from two given coordinates . . . As we advance in one cost dimension we must know how far we have gone in the other two, so that the marginal cost of any one *product* can be calculated only if we assume as given both the number and quantity of all other products.[21]

While the Commission's cost studies have not explicitly analyzed variability

[21] J. A. McDonald, "Some Notes on the Economics of Transportation," *Canadian Journal of Economics of Political Science*, November 1951, pp. 517-18.

with respect to weight and distance separately nonetheless the several motor carrier regional studies contain data which permit such an assessment.

Tables I and II show the behavior of costs as weight changes for any given distance, as distance changes for any of the given weights and for changes in both weight and distance. Let us explore these relationships further. For any particular distance, total cost per shipment rises less than proportionately as

TABLE I—Class I Motor Common Carriers of General Commodities, Eastern-Central Territory, Out-of-Pocket Cost for Varying Combinations of Weight and Distance: 1956

(dollars per shipment)

1000 mi.......	4.88	7.14	9.32	15.61	33.39	66.74	134.80	265.57	373.02	497.45
900 mi.......	4.74	6.85	8.89	14.76	31.35	62.29	124.81	243.72	339.34	455.35
800 mi.......	4.59	6.57	8.46	13.93	29.31	57.84	114.83	220.19	305.67	409.44
700 mi.......	4.46	6.29	8.05	13.10	27.27	53.71	105.55	198.33	271.99	363.52
600 mi.......	4.32	6.03	7.65	12.32	25.37	49.26	96.78	178.17	238.32	321.43
500 mi.......	4.19	5.76	7.25	11.48	23.47	45.13	86.30	156.31	207.23	279.34
400 mi.......	4.06	5.49	6.85	10.70	21.43	41.00	77.03	134.46	176.15	233.42
300 mi.......	3.92	5.22	6.44	9.87	19.54	36.55	67.75	112.61	145.06	191.33
200 mi.......	3.78	4.94	6.03	9.03	17.50	32.10	57.77	90.76	113.98	145.41
100 mi.......	3.65	4.69	5.64	8.31	15.60	28.29	49.21	68.91	82.89	103.32
50 mi.......	3.57	4.52	5.39	7.77	14.43	25.42	42.79	55.47	64.76	80.36
Weight (lbs.)..	100	200	300	598	1,458	3,178	7,132	16,808	25,904	38,265

Source: Data derived from *Cost of Transporting Freight by Class I Motor Common Carriers of General Commodities, Eastern-Central Territory, 1956,* Bureau of Accounts, Cost Finding, and Valuation, Interstate Commerce Commission, Statement No. 4-57, Washington, D. C.., October 1957.

weight increases. This is seen immediately in Table II where average cost per ton-mile declines as weight increases. Furthermore, the rate of increase of cost with respect to weight rises as weight increases. For example, the percent variability of costs per ton-mile with respect to weight (holding distance constant at 600 miles) for a weight increase from 200 pounds to 300 pounds is 54 percent whereas, taking a weight increase from 1,458 pounds to 3,178 pounds for a 600-mile shipment, yields a variability of 80 percent. We may

TABLE II—Class I Motor Common Carriers of General Commodities Eastern-Central Territory, Marginal Cost per Ton-Mile for Varying Combinations of Weights and Distance: 1956

(dollars per ton-mile)

1000 mi.......	.098	.072	.062	.052	.046	.042	.038	.032	.029	.026
900 mi.......	.105	.076	.065	.055	.048	.044	.039	.032	.029	.026
800 mi.......	.115	.082	.071	.058	.050	.046	.040	.033	.030	.027
700 mi.......	.127	.090	.077	.062	.053	.048	.042	.034	.030	.027
600 mi.......	.144	.101	.085	.068	.058	.052	.045	.035	.031	.028
500 mi.......	.168	.115	.097	.077	.064	.057	.049	.037	.032	.029
400 mi.......	.203	.138	.114	.089	.073	.064	.053	.040	.034	.030
300 mi.......	.261	.174	.143	.110	.090	.077	.063	.045	.037	.033
200 mi.......	.378	.247	.202	.151	.120	.101	.081	.054	.044	.038
100 mi.......	.730	.469	.376	.277	.214	.178	.138	.082	.064	.054
50 mi.......	1.428	.905	.719	.518	.395	.322	.240	.132	.100	.084
Weight (lbs.)....	100	200	300	598	1,458	3,178	7,132	16,808	25,904	38,265

Source: Data derived from *Cost of Transporting Freight by Class I Motor Common Carriers of General Commodities, Eastern-Central Territory, 1956,* Bureau of Accounts, Cost Finding, and Valuation, Interstate Commerce Commission, Statement No. 4-57, Washington, D. C., October 1957.

therefore conclude that costs rise slowly with respect to weight increases from low levels but the rate of increase rises as weight changes from successively higher levels. This of course means that costs per ton-mile decline sharply at first as tons increase but the rate of decline tapers off markedly and approaches zero for heavier weights.

For any particular weight, total cost per shipment likewise rises less than proportionately with respect to increases in distance. There is a similar pattern of cost behavior with respect to distance (ceteris paribus) as found above with respect to weight. For example, the variability of costs with respect to ton-miles as distance increases from 100 miles to 200 miles (holding weight constant at 1458) is 12 percent whereas the variability for a distance increment from 600 miles to 700 miles is 42 percent. We may similarly conclude that costs rise slowly with distance as length of haul increases from low levels but the rate of increase rises as distance increases from higher levels. With both distance and weight increments, however, the increase in cost is less than proportionate. Consequently, we find that average cost per ton-mile per shipment declines as either (or both) tons and miles increase.

To which variable is cost more sensitive? It is evident that costs are more sensitive to weight changes than to varying lengths of haul. However, it is nonetheless true that cost variability depends upon the magnitudes of both weight and distance (holding one constant and allowing the other to vary) with which one starts. That is to say, the percent variability of weight (or distance) is greater the greater the (constant) distance (or weight) which one uses. For example, the variability of cost with respect to a weight change from 598 pounds to 1458 pounds for a given distance of 50

miles is 60 percent whereas if the given distance is 1000 miles the variability is 80 percent. This, of course, reflects the fact that the degree of decrease in average cost per ton-mile tapers off with increases in either tons or miles.

This indicates that the extent to which weight or distance affects costs depends upon the level of either variable with which one starts. However, taking the *average* weight and distance found in the Interstate Commerce Commission study we can compute the variability of costs with respect to both weight and distance. This will indicate something of a "normal" or at least average situation and will permit us to say something definite about the relative impacts of weight and distance upon costs. Admittedly, since the degree of dispersion about the mean is unknown, the extent to which the average is truly representative cannot be known. With this reservation, we may examine weight and distance changes from the given averages.

Table 8 of the Interstate Commerce Commission study[22] gives an average haul for the carriers studied of 570 miles while Table 10 gives the average weight per shipment as 1,628 pounds. Taking the closest approximation to these figures from Tables I and II above, we may assume an average haul of 600 miles and an average weight of 1458 pounds. Using these as points of departure, the variability of costs with respect to distance is 42 percent while the variability with respect to weight is 80 percent. If this represents the "normal" situation we can then suggest unequivocally that cost is more sensitive to weight than to distance. This, of course, means that costs per ton-mile will decline more rapidly as the proportion of miles to tons arises. Thus, even in motor carrier transportation there appear to be distinct economies up

22 I.C.C. Statement No. 4-57, *op. cit.*, p. 16.

to at least 1000 miles after which point the degree of cost taper appears negligible. This finding is consistent with that of Roberts who concluded, after examining 114 Class I carriers of general commodities in Central Territory, that the influence of average haul in reducing costs per mile was greater than the influence of route utilization.[23] The different costs for the identical ton-mile value having different proportions of tons and miles can be seen readily if we take a 50-mile shipment of 200 pounds. This gives a cost of $4.52. Whereas a 100 mile shipment of 100 pounds (which is the same in terms of ton-miles as 200 pounds shipped 50

[23] Merrill J. Roberts, "Some Aspects of Motor Carrier Costs: Firm Size, Efficiency, and Financial Health," *Land Economics*, August 1956, pp. 232-3.

miles) costs only $3.65. This also bears out our above conclusion that costs are less sensitive to distance than to weight.

Graphically, Table I would appear as shown in Figure 1. The whole of Table I has not been drawn, only the set of cost values for distances up to 600 miles and weights up to 3178 pounds. The analogy between this and the production surface of economic theory is obvious. However, the production surface as usually drawn is symmetrical and assumes that the production function is linear and homogeneous of the first degree. In the present instance we have already noted the lack of symmetry and indeed Figure 1 shows costs rising more rapidly with weight than with distance. It is also readily seen that equal ton-miles with

ILLUSTRATION OF COST BEHAVIOR IN TRANSPORT
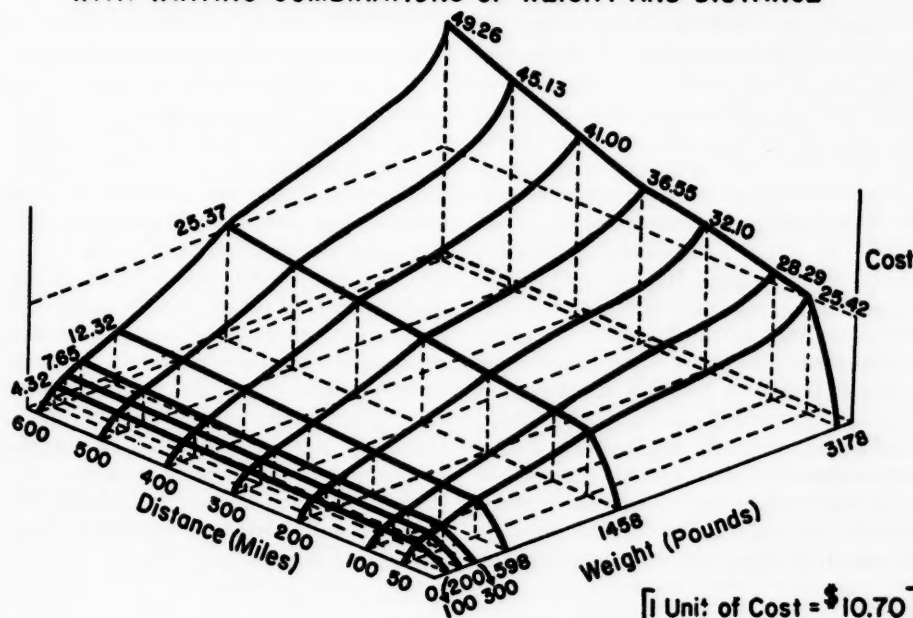## WITH VARYING COMBINATIONS OF WEIGHT AND DISTANCE

Figure (I)

Data Taken From Table I

| 1 Unit of Cost = $10.70 |
| Scale: 1 Unit = ½ Inch |

varying proportions of tons and miles will have different costs.

Only by holding one variable constant (along with velocity) is the homogeneity assumption on the cost side valid. However, even here it is not very relevant. In production theory, the variable items are input items (usually identified as land and labor or capital and labor) and may freely be varied at the discretion of the firm in an attempt to minimize costs. However, while in transportation both weight and distance may be construed as input items (i.e., *loading* inputs and *moving* or *line-haul* inputs) producing a product called ton-miles, they are *not* freely variable by the producing unit. The varying proportions of tons and miles are primarily (though not entirely) a function of buyers' (i.e., shippers') policy as manifest by location, packaging, rate of output as well as inventory policy. Only to a very minor extent can a transportation firm influence these items by itself. It is therefore reasonable to suggest that the transportation enterprise has practically no discretion over the proportion of loading and moving inputs either for a specific shipment or an aggregate of shipments. The primary concern of a transportation enterprise is to attempt to minimize the cost of loading and moving inputs (taking each separately) taking the proportion between the two as given by and large. It is, of course, true that where a particular shipment is less than either a truckload or carload, the ratio of tons to miles actually performed is partly a function of the utilization of the truck or box car. However, since rates are computed between specific points or for actual mileage blocks, it is presumed that generally the

shortest distance will be used. Only within this relatively limited range does the transport firm have influence over the proportion of tons and miles. Even here, the influence relates primarily to distance although carrier discretion exists with respect to the amount of l.c.l. or l.t.l. shipments per running unit.

Generally, then, the heterogeneity on the cost side is established by three features: (1) the identity of costs for various amounts of ton-miles, (2) the variety of costs for the same total ton-mile figure (costs and miles being in different proportions) and (3) the substantial lack of discretion on the part of the transport firm over the proportions of weight and distance. All of these render production theory much less applicable to transportation and support the view of those who have stressed the multi-product aspect of transportation enterprises.

Thus the unit of output in transportation is a rather elusive animal. The foregoing analysis also suggests that there is no cost curve in the economic sense for the total output of any transportation enterprise. There are a variety of costs coexisting at a moment of time, depending upon specific conditions. The variability of the output unit is perhaps the most important factor complicating cost finding in transport. Analysis of the economics of transport would become more realistic if the above complexities were given more careful consideration. Certainly any presumption of output homogeneity does not advance our knowledge of the transportation field nor make it any easier to obtain a cost-based rate structure.

Property Rights, Tenancy Laws of Cuba, and Economic Power of Renters

Nature of Property Rights

PERHAPS the most ancient problem still plaguing mankind in every corner of the world is the terms on which he uses land for agricultural purposes. Recent events in Cuba indicating a political, economic, and social unrest partially related to land use conflicts attest to the continuing significance of this problem. The expanding revolutionary force over the past couple of years in Cuba gained its first strength in the fertile sugar-growing, but heavily squatter-share-cropper-renter-populated Province of Oriente —and returned there to launch its land distribution program. Such a sequence is not the first in the history of the Latin American Republics.

The tiller has doggedly sought the "right" to work the soil increasingly on "his" terms and for "his" reward. It mattered little whether he was a roaming hunter in the primitive era or a cotton farmer facing an allotment in the twentieth century; the conditions of land use were his major vexations. Today, with attention focused upon problems of economic development, the ratio of population to land area receives star billing. However, closer scrutiny frequently produces the substantive problem; namely, the particular structure of relationships affecting the farmer as he with changing technology uses his land area, whether small or expansive.[1] A problem consistently facing the land user, though of course not the only one, is basically that of the nature of his property rights.[2]

During the economic stages of the feudal states and early autocratic empires the worker on the land had little control over the soil he tilled or the produce resulting from his efforts. In fact, he became a form of capital for the owner of the land, clearly exemplified in the periods of slavery in this country and others. The worker first thought he had to secure uncontested control over the use of his land if his efforts were to result in an enhanced level of living. His goal, in those countries exhibiting land control as a prime means of private power, was some form of property ownership which would be accessible to him and be guaranteed by governmental and legal authority. The culmination was such forms of ownership as primogeniture or more recently fee simple. Often this unbending desire for land ownership has resulted in the parcellation of land holdings exemplified by the minifundios in Latin America.

However, the goal of ownership of all the resources needed for farm production received a second thought when economical operation began demanding increasing amounts of capital. Furthermore, the farmer, much as his urban brother in a highly industrial economy, had begun to strive for improved economic power through means other than just land ownership—such as seeking stable markets, improved research, bargaining power or organizations, governmental sanction of production control, and even retirement security. Economic power here refers to the capacity of the individual to achieve deliberately improved economic returns or standards of living.[3] As he has

[1] Testimony from every corner of the world attesting to the significance of these relationships is reproduced in the following volume: K. H. Parsons, R. J. Penn, P. M. Raup, eds., Land Tenure; Proceedings of the International Conference on Land Tenure and Related Problems in World Agriculture (Madison: University of Wisconsin Press, 1956).

[2] John R. Commons traces this struggle through the centuries of history in Western Europe where our concept of "rights" took form and records it in his Legal Foundations of Capitalism (New York: The Macmillan Company, 1924), pp. 214 f.

[3] Such a conceptualization of economic power is an expansion of the meaning often used in the history of economic thought. For a more adequate discussion of the concept, see the unpublished Ph.D. thesis of the senior author entitled, An Analysis of the Act as an Economic Phenomenon; the Concept of Economic Power and the Problems of Agriculture, University of Wisconsin, 1954.

sought improved living via these additional rights, freedoms, or opportunities, his quest for personal and complete control over the land he works has seemed to diminish. In fact, farmers have seemed to fight for such rights as a "fair" market price, available electric power, water supply, and "reasonable" credit almost as hard as they have for exclusive land use rights.

Just as important to the farmer as the right of individual land ownership was the simultaneous struggle for public control of the government which guarantees his rights. Once such a government, whose policies were more "public" than "private," was achieved, the people have seemed to become interested in "rights" in addition to that of land ownership.[4] The individual's only protection was then not just in the undisputed holding of property, but in the using of it in such a manner as to improve his livelihood through a myriad of exchanges and activities of a complex economy. Truly, property has come to mean a "bundle of rights" protected by and developed in the processes of a publicly constituted government.[5]

Some economic investigators have suggested, and offered some proof, that farmers with limited capital could secure greater economic returns by borrowing someone else's land as renters and use the scarce, owned resources in the form of livestock, machinery, and operating capital.[6] A serious deterrent to many farmers pursuing this alternative, particularly in the South, is the uncertainty related to both the duration of a rental arrangement and the security of any value added to the farm by the user. In fact, the rights of the tenant in the United States essentially are limited to those he can bargain and specifically contract for in his negotiations with the land owner.[7] Uncertainty thus

associated with tenancy can contribute to inefficiency in resource utilization the same as to small operator units when owner-operatorship is carried to the extreme.

Purpose of Inquiry and the Cuban Example

Regardless of the similarity of the struggle of the farmers in various countries to control their land or of people to control their governments, the precise structure of the property rights that emerge may show considerable divergence. These rights apparently need to be patterned uniquely to blend into the particular man-land ratio, stage of industrialization, structure of existing rights, customs of the people, and social value systems.[8] Hence, the systems of property rights may not be transferred with dependable success from one country to another. Similar problems also may be encountered in transplanting other practices and policies in the process of economic development.

Does this mean that nothing can be gained from examining the property rights of other countries? Surely not. One theory of knowledge holds that theoretical constructs function primarily to help establish alternative courses of action or modes of response, not to posit what should be done.[9] Hence, if alternative property rights and their results in different countries can be examined, ideas might be gleaned that are relevant to the problems facing other societies.[10] This inquiry is pursued for such a purpose.[11] An

[4] "Rights" are so often spoken of as being given to the people by some original document or government itself. Such expressions perhaps have more truth when a dictatorial form of government exists than in other forms. In some form of representative government, "rights" are sought and won by the quest and compromise of many private interests who see the gains as superior even in the face of new duties which always accompany the rights. Such processes are perhaps the best means of determining the public value and are the source of basic public policy. *Ibid.*

[5] Richard T. Ely and G. S. Wehrwein, *Land Economics* (New York: The Macmillan Company, 1940), p. 76.

[6] For example, T. R. Schultz, "Capital Rationing, Uncertainty, and Farm Tenancy Reform," *Journal of Political Economy*, December 1940, 309-324.

[7] A committee of rural social scientists concluded at a conference on Family Farm Policy in 1946: "For the great majority of tenant farms, however, significant and rapid

improvement can be brought about only by legislation and official regulation of landlord-tenant relations. In most areas the bargaining power of the landlord is so much greater than that of the tenant that the tenant cannot bargain effectively even if he is fully aware of the defects in the leasing arrangements." Joseph Ackerman and Marshall Harris, eds., *Family Farm Policy* (Chicago: University of Chicago Press, 1947), pp. 488.

[8] K. H. Parsons, "Economic Citizenship of the Land," *Land Economics*, February 1954, pp. 26-28.

[9] Such a theory is elaborated by K. H. Parsons in his "Logical Foundations of Economic Research," *Journal of Farm Economics*, November 1949, particularly sections on pp. 671-81.

[10] For example, the Agricultural Holdings Acts of Great Britain have been studied for almost a century as a classic system of tenancy rights. For a recent discussion see a presentation made at the International Conference on Land Tenure and Related Problems in World Agriculture. Parsons, Penn, Raup, *op. cit.*, pp. 366 f.

[11] For a summary of an almost exhaustive and certainly useful analysis of completed research in the United States on property and, particularly, tenure relations, see Leonard S. Salter, "Research in Landed Property," *A Critical Review of Research in Land Economics* (Minneapolis: University of Minnesota Press, 1948), chap. vii.

attempt is made to analyze some of the effects which the tenancy laws of Cuba, unique to those of many other American countries, may have had upon the relative economic power of the "renting" and "owning" farm operators.

The empirical evidence examined in this investigation was obtained from a sample of farms taken during 1953 in an area of Western Cuba near Havana City where the family farm predominates.[12] 'In this area the project leaders concluded that the soils, products, systems of farming, and economic land classification were relatively uniform.

In Cuba, as is common throughout Latin America, agriculture remains the dominant industry. One-half of the agricultural land was used for production of sugar cane in 1952.[13] Furthermore, in 1946 about seventy percent of the total farm operators could be classed as tenants.[14] When rights emanating as public policy from some form of representative government are under consideration, the nature of the private interests participating in the policy development is of importance. The renters of Cuba at the time of this study had developed a Colonos Association with some sixty thousand (60,000) active members.[15] Its influence in governmental action appeared to be considerable, both indirect and via some of its members who sat in the federal legislative bodies. In addition to the influence which this interest group has had upon the land use policies of Cuba, that country has in recent decades had a vital interest in fostering and protecting the industry that produced its prime export product—sugar.

Structure of Cuban Tenancy Laws

Contemporary property laws in any country are almost invariably influenced by prior historical occurrences and statutes. Such happenings during the development of Cuba may tend to illuminate its subsequent tenancy laws. For several hundred years, beginning in the sixteenth century, the system of large land estates (*encomiendas*) was instituted via grants of land by the Spanish rulers to religious bodies (for the purpose of converting the natives) and to landowners (for purposes of developing huge livestock farms). The process of division of such estates by way of the claims of succeeding generations of heirs resulted in the *hacienda comunera* (farms held in common).

No sooner had the evolutionary system of family type farms, held in common, been dissolved by government decree than a trend was reversed toward larger land holdings (*latifundia*) around the turn of the eighteenth century by the emergence of the sugar industry. The profitability of the sugar enterprise in this new industry at the doorstep of the rapidly expanding United States economy attracted both Cuban and foreign capital to the sugar plantations. Thus, today, the majority of the Cuban farmers find themselves in the status of tenants.[16]

Subsequent to the maturation of the Cuban sugar industry as a producer of a major export commodity, trade difficulties, excessive production, and economic crises abroad prompted the enactment of many governmental policies vital to this segment of the economy. Some of these curtailed production via allotments. Quotas were applied to the exports, at times in combination with importing countries; minimum wages were established for agricultural labor; maximum rents were set forth; producers were granted concessions in credit use; maximum sizes of land holdings were established; the government granted rights to renters on private lands not being used for appropriate production.[17] During the same period, two sets of tenant rights of significance to this inquiry were established: "action rights" and "permanency rights."[18]

[12] The sample area was a part of the Organization of the American States' Northern Zone demonstration unit at that time, located in the Barrios of Bejucal, Quivican, and La Salud, Cuba.

[13] *Anuario Azucarero de Cuba,* 1952.

[14] *Censo Agricola de Cuba de 1946.*

[15] A colono is a person who rents land from the sugar mill owner with the objective of producing sugar cane, but the Colonos organization also includes those renting from individual land owners. The renter normally makes the managerial decisions, except when he rents from a mill operator, in which case the latter decides how much sugar cane will be planted and on what acreage.

[16] Lowry Nelson has observed: ". . . . it is the point of view of this work that in spite of certain modifications, there are vestigial remains of the feudal class system still subtly persisting in Cuban society . . . ." *Rural Cuba* (Minneapolis: Colwell Press, 1950), p. 143. For a careful characterization of the evolution of the current Cuban land system, see Chapter V.

[17] For a study of the Cuban domestic policies affecting the sugar industry, primarily those related to wages, pricing, and marketing, see B. C. Swerling, "Domestic Control of an Export Industry: Cuban Sugar," *Journal of Farm Economics,* August 1951, pp. 346-56.

[18] The actual terms used during the legal and institutional development of these rights varied over the 20-to 30-year period. Action and permanency, the terms appearing in this analysis, are among the English counterparts found used in referring to the rights, which in Spanish often appear as *accion* and *tenencia.*

A. *Action Rights.* Cuban renters have rights, guaranteed as are property rights by law, in the sugar cane roots planted and in other improvements made during their tenure. These rights are referred to as action rights and have emerged gradually during the period from 1922 to 1948.[19] In the case of sugar production, the cane roots once planted can produce stems for a period of six years. Planting of cane roots by the tenant thus tends to compare with constructing a building, erecting a fence, or improving the soil.

How does the action right enter into negotiations for the sale of other accompanying property rights? The tenant can either sell his action rights to another renter approved by the landowner or to the owner himself. The value of the action right appears to be determined by factors similar to those affecting land values, such as product price, scarcity of the property right, and productivity potential. Action rights were found in the study area to equal the value of the land upon which the tenant acquired his right in a majority of the cases. The average action right value for the forty-six tenants, according to their estimates, was $8,000.

Sub-renting is not permitted and the laws specify that rental contracts must be written. However, among the sample of renters interviewed in this study, only one-third possessed a recorded contract. The custom seems to call only for a receipt of rental payment.

B. *Permanency Rights.* A second significant group of rights was acquired by the renter via Cuban legislation enacted from 1937 through 1952—the permanency rights.[20] These insure the tenant the right to continue using the land he occupies as long as he fulfills two conditions: first, production of his sugar cane quota (assigned by law to the operator or producer); second, payment of rent at the appropriate time. The rent is normally determined by law for sugar production (utilizing a percentage of the total yield or a maximum percentage of the sale price for the land being rented) and by negotiation between the landlord and renter for other enterprises.[21]

Those regulations establishing the quotas and rentals seem to favor both the smaller producer and renter.

In the area included in this study the landlords were found not to be residents on the rented farms but usually lived in cities. Some had acquired the land as a business venture while others became owners through inheritance. As might be anticipated in the presence of such property laws as those insuring the permanency right, the renters stayed on a particular farm for considerable periods of time. In the study area some farms had been rented by the same families for more than one generation.

Economic Status of Renters and Owners

Cuba has one of the higher per capita incomes in Latin America, yet this tells little about the distribution among farmers.[22] The concentration of ownership of land in the hands of one segment of the farmers of Cuba is still much in evidence. Thirty-six percent of the land area in farms is found in half a percent of the total number of farms; 3 percent of the area is in 40 percent of the farms (1946).[23] In 1946 there were still 114 farms each with 24,700 acres of land and the larger farms tended to predominate in the eastern provinces.[24] Furthermore, as the sugar industry has developed, its relative importance as a farm product has shifted somewhat toward the eastern portion of the country where recently the provinces of Camaguey and Oriente accounted for half the total production.[25]

One of the important considerations of any policy or system of property rights is the effect it has upon the economic sectors to which it is directed. In this case, what is the effect of the action and permanency rights upon the relative economic gains of the renters and owner-operators? This study did not permit an actual determination of the economic position of the renters as being causally related to the tenure laws; rather, the economic characteristics were simply found to exist

[19] For example, Land Lease and Sharecropping Law of 1948.

[20] Established by Sugar Coordination Law, Article 26; Colonos Statutes, Article 97; and Law No. 249 of March 1952.

[21] Somewhat similar rental policies were reported for Brazil by Joao Goncalves de Souza in, Ackerman and Harris, *op. cit.*, pp 251 f.

[22] Cuba's per capita income was exceeded in South America only by Argentina, Uruguay, and Venezuela according to the United Nations Publication, *National and Per Capita Incomes in Seventy Countries, 1949*, 1950, p. 28.

[23] *Report on Cuba; Findings and Recommendations of an Economic and Technical Mission Organized by International Bank for Reconstruction and Development and Government of Cuba 1950* (Baltimore, Maryland: John Hopkins Press, 1951), p. 80.

[24] Nelson, *op. cit.*, p. 135.

[25] *Report on Cuba, op. cit.*, p. 796.

TABLE I—Total Acreage per Farm and Total Acreage Under Cultivation per Farm, Sample of Farms in Havana Province, Cuba: 1952*

Type of Farm	Number of Farms	Total acreage— Mean (Besanas)	Total acreage in cultivation— Mean (Besanas)	Results of 't' test for significance of difference be- tween Means
Small family farm-owner..........	12	57.3	42.9	Not significant
Small family farm-renter..........	31	53.8	44.7	
Large family farm-owner..........	14	84.6	65.3	Significant
Large family farm-renter..........	11	110.2	80.7	
Multi-family farm-owner..........	7	335.1	256.1	Not tested
Multi-family farm-renter..........	4	260.0	252.0	
Total farm-owners...............	33	133.2	100.2	Significant
Total farm-renters...............	46	86.4	71.4	

* One acre equals 1.63 besanas.

TABLE II—Total Man Equivalents per Farm, Sample of Farms in Havana Province, Cuba: 1952

Type of Farm	Number of Farms	Mean	Results of 't' test for significance of differ- ence between Means
Small family farm-owner......................	12	1.6	Significant
Small family farm-renter......................	31	1.8	
Large family farm-owner......................	14	3.1	Significant
Large family farm-renter......................	11	4.2	
Multi-family farm-owner......................	7	16.8	Not tested
Multi-family farm-renter......................	4	11.7	
Total farm-owners............................	33	5.46	Significant
Total farm-renters............................	46	3.25	

after these rights had been in operation for between five and thirty years.

Relative economic well-being can be ascertained in a multitude of ways. The following comparisons utilized in this inquiry are frequently used and serve as important indications of both consumption and productive capacity: size of farm, area under actual cultivation per farm, total man equivalents per farm, total capital per farm, gross farm income per man equivalent.

From the study area in Cuba, 79 farms were selected for the sample.[26] Thirty-three of these farms were owner-operated (100% sample) and 46 were renter-operated (50% random sample). For the subsequent analyses, the farms were sub-divided into three

[26] For elaboration on the organization of the study see the unpublished Master's thesis of the junior author entitled, *The Cuban Land Tenure Laws and Some of Their Economic Effects upon Sugar Production*, University of Tennessee, 1954.

TABLE III—Total Capital per Farm, Sample of Farms in Havana Province, Cuba: 1952*

Type of Farm	Number of Farms	Mean	Results of 't' test for significance of difference between Means
Small family farm-owner......................	12	$15,140	Significant
Small family farm-renter......................	31	8,410	
Large family farm-owner......................	14	17,220	Not significant
Large family farm-renter......................	11	17,870	
Multi-family farm-owner......................	7	88,191	Not tested
Multi-family farm-renter......................	4	65,325	
Total farm-owners...........................	33	33,550	Significant
Total farm-renters...........................	46	15,950	

* Capital includes land and buildings, machinery and tools, irrigation equipment, work animals, and livestock.

TABLE IV—Gross Farm Income per Man Equivalents, Sample of Farms in Havana Province, Cuba: 1952*

Type of Farm	Number of Farms	Mean	Results of 't' test for significance of difference between Means
Small family farm-owner......................	12	$1,440	Not significant
Small family farm-renter......................	31	1,243	
Large family farm-owner......................	14	1,046	Significant
Large family farm-renter......................	11	1,566	
Multi-family farm-owner......................	7	1,794	Not tested
Multi-family farm-renter......................	4	1,908	
Total farm-owners...........................	33	1,350	Not significant
Total farm-renters...........................	46	1,381	

* Gross farm income includes cash receipts and an estimate for home consumption of farm production.

groups according to man equivalents per farm: (1) small family farms, less than 2.5 family man equivalents, (2) large family farms, 2.5 to 4.9 family man equivalents, (3) multi-family farms, 5.0 or more family man equivalents.[27]

[27] One man equivalent was considered a farm person, 18 years or older, working on the farm approximately an 8-hour day during the work days throughout the year. The farms were family farms in that they provided essentially full-time employment for the families living on them and such families provided most of the labor used.

Comparisons for 1952 between the renters and owners in the selected areas in Western Cuba for each of the indicators mentioned earlier as measures of economic well-being appear in Tables I-IV.

Summary

The quest for a "desirable" system of allocating rights to capital between the provider and the user, between the person offering the objects of production and the persons

involved in the production will undoubtedly continue. It could well be said that much of the history of agriculture has been overshadowed by the search for a satisfactory relationship between renter and owner.[28] Property rights are the focal point in this inquiry.

Experiments involving various combinations of property rights are in process throughout the world as countries wrestle with their land-use problems, and may provide useful ideas and data for the social scientist studying land tenure. The findings from these experiments can become increments of knowledge that may be utilized in examining alternative tenure arrangements in any country. To ascertain such findings was the purpose of this inquiry into the nature of tenancy laws of Cuba and the relative economic status of the owner-operators and renters five to thirty years subsequent to the enactment of the laws in a small study area near Havana.

Land tenure problems certainly still persist in this Caribbean country. However, the recent armed political revolt in Cuba gained its original popular support partially from the distress of landless farmers in heavily populated, eastern-most Oriente Province; there, sugar is a dominant product, large farms exist, and perhaps most importantly a fifth of the farms are run by squatters.[29]

The tenancy laws of primary concern in this study were: (1) the action right, which guarantees the renter reimbursement for his improvements added to the farm, and (2) the permanency right, which guarantees the renter occupancy on the farm upon the conditions of rent payment and production of his sugar quota.

To assess the relative economic characteristics of the renter and the owner-operator classes of farmers, the five following comparisons were made empirically using data obtained in the Western Cuban Province of Havana with the stated results:

(1) *Total acreage per farm.* Except for the large family farm, the renter was found to be operating on somewhat fewer acres than was the owner-operator.

(2) *Total acreage under cultivation per farm.* Even though the total acreage per farm in (1) above was generally larger for the owner, the renter was found to be tilling generally as much acreage as the owner. Hence, the advantage which the owners could have by operating larger farms is somewhat reduced with the smaller proportion of the farm being tilled.

(3) *Total man equivalents per farm.* The renter was found to be using slightly more labor and a greater degree of intensification on his smaller acreage than was the owner except on the large multi-family type of farm.

(4) *Total capital per farm.* Considerably more capital was being used by the owner group on the small farms while the renters' capital compared somewhat more favorably to the owners' on larger units.

(5) *Gross farm income per man equivalent.* Even though the renter was utilizing slightly less land and somewhat less capital than the owner-operator, the gross farm income to the renter per man was generally comparable to that achieved by the owner.

The foregoing empirical findings cannot be determined as causally related to the tenancy laws examined herein. Rather, the conditions were found at a time when such tenant rights as action and permanency had been in effect for a period of from five to thirty years. It would, however, be strongly suggested that these rights, along with many other concessions granted the renter in Cuba such as maximum rents, credit assistance, and minimum sugar quotas, have resulted in some additional economic power accruing to the farmer who operates a farm as a renter. Political, economic, and social upheavals, such as that experienced recently in Cuba, may well be related to the terms on which the tiller of the soil uses his land; the unique rights examined in this study and may other public measures might need to be considered as improved tenure arrangements are evolved.

R. G. F. Spitze

Professor of Agricultural Economics
University of Tennessee

Gregorio Alfaro A.

Agricultural Economist,
American Embassy,
San Jose, Costa Rica

[28] K. H. Parsons, "Land Reform in the Postwar Era," *Land Economics,* August 1957, pp. 213-227.
[29] For recent reports on these developments, see H. L. Matthews, "A New Chapter Opens in Latin America," *New York Times Magazine,* January 11, 1959, pp. 13, 72-4, and Fidel Castro, "What Cuba's Rebels Want," *Nation,* November 23, 1957, pp. 399-401. For a concise account of the background of the current uprising, see H. L. Matthews, "Cuba in Ferment," *Foreign Policy Bulletin* (New York: Foreign Policy Association, April 1 ,1957), pp. 109-12.

A Study in Excess Capacity

THE lack of attention paid to local monopolies in economic analysis is curious. It is true that large firms are important and have a profound influence on our lives. But perhaps the omnipresent local markets, where attributes of monopoly are not always apparent, have at least as much influence. United States Steel may have more difficulty in covering up monopolistic practices than John Jones, druggist.

The theory of monopoly is easily subject to condemnation when it is oversimplified. To meet this criticism complications may be added but this in turn invests the theory with less generality. Another alternative is to find strong cases. The parking lot industry seems to be such a case. The product is simple and the inputs are unskilled labor and relatively unimproved land. The only significant factor distinguishing one lot from another is its location.

The usual earmarks of monopoly are relatively high prices coupled with excess capacity. A statistical measure of excess capacity and its relationship to monopoly, utilizing data from the center-city Philadelphia parking market follows.

Nature of the Data

The data were copied from the records of the Philadelphia tax office. The area used was bounded by the Schuylkill and Delaware Rivers and South and Vine Streets as shown in Figure I. The center-city was divided into twelve market areas.[1]

The success of this zoning scheme is evident from the fact that in ten analyses of variance for this zoning scheme (an analysis of variance of prices and quantities demanded for the different classes of demand such as "one-hour," "two-hour") the F-ratio does not exceed the one percent level of significance except in the case of quantities demanded for "evening" parking. In the latter case the F-ratio did not exceed the five percent level of significance. Thus, for example, 109 lots were examined in 11 zones (Zone 5 had no "all-day" parking) in regard to the number of cars parked "all-day." The "between-zone" variance was 84.63 and the "within-zone" variance was 17.94. The F-ratio was 4.72 which did not exceed the one percent level of significance.

Table I shows the concentration ratios for each market area. The concentration ratios compare the capacity of the larger firms to the total capacity of the market area. It will be noted that the market areas adjacent to City Hall at Broad and Market Streets generally have a higher concentration ratio than the market areas on the fringe of the center-city area. These concentration ratios do not suffer the fault of the census classifications because the firms are all producing a related product and all competitive lots are grouped into market areas.

Since the areas near the center contained the highest concentration ratios, the entire center-city area was further classified into two areas termed the "high concentration" area and the "low concentration" area. The shaded portion of Figure I represents the area of high concentration. In this area one firm controls 63.2 percent of all capacity. There were 30 lots in the high concentration area and 40 lots in the low concentration area.[2]

There were no municipal lots. Furthermore, with one major exception at 8th and Market Streets, practically all "free" or "reduced-rate" parking for shoppers offered by businesses was operated by the commercial lots. This arrangement, it should be noted, may result in price discrimination in favor of the shopper.

Street parking was of no significance. Specifically, at the beginning of March 1953 daytime street parking in the center-city was generally banned. This increased the importance of off-street lots by eliminating a near-substitute.

[1] For further discussion of this zoning scheme see the writer's "Product Differentiation in the Parking Market," *Land Economics*, August 1958, pp. 247-249.

[2] There were 228 lots located in central Philadelphia during January 1955. The 70 lots included in the data are the more important lots, with larger than average capacities. Many of the excluded lots were "monthly" lots, offering space at monthly prices. Since the lots which were included in the analysis had a greater capacity than did the excluded lots, the total spaces included in the analysis is much higher than the proportion of lots indicates.
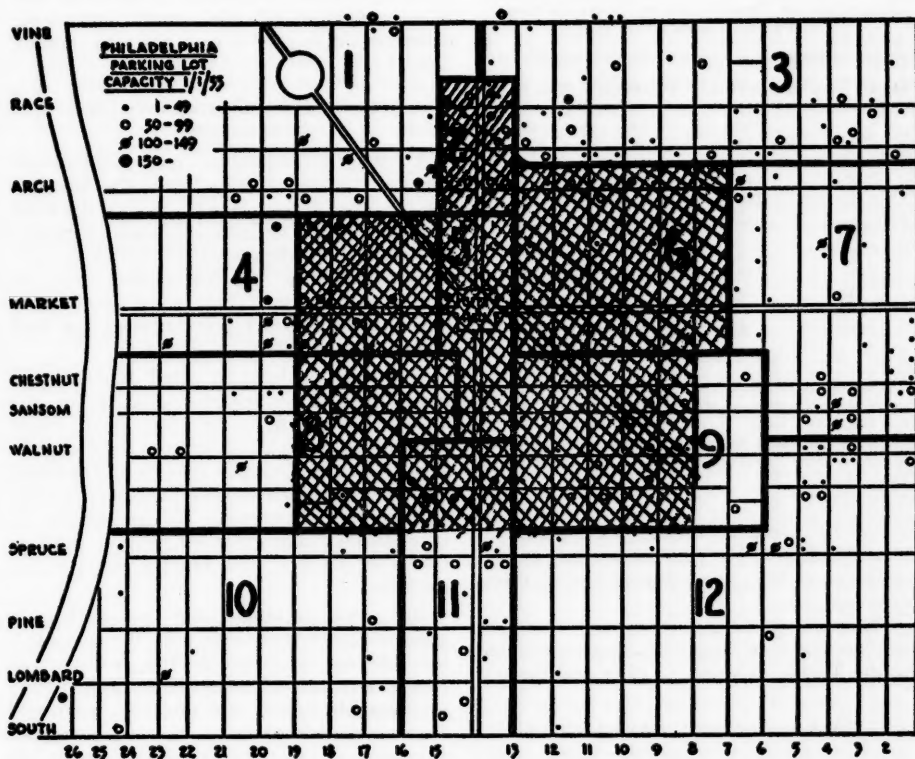
Capacity of lots was determined by the Traffic Engineer's Office. Capacity utilization was based on the tax investigator's inspection of lots. The information was acquired by city investigators who made personal visits to the 70 lots during the month of March 1953. The investigators visited each lot once a day, excluding week-ends and March 1st and 2nd, there being nineteen daytime weekday observations for each lot during the month. The investigators recorded the time of day and the number of cars on the lot. In general, the time of inspection appeared to be spread over the day with some concentration around one and two o'clock.

Absolute Excess Capacity

The concept of excess capacity implies that there is some benchmark of full capacity.

Since the cost function is probably linear,[3] the economic (minimum cost) and the Traffic Engineer's concept (maximum number of parkers after allowing for fire and aisle space) coincide.

Table II illustrates the fact that lots usually operate at less than full capacity during a large portion of the day. The third column shows for each market area the mean value of the ratio of the number of cars parked to the capacity of the lot. For the entire center-city area this is around 85 percent. The fourth column indicates for each market area the mean value of the ratio of the greatest number of cars observed parked to the capacity of the lot. This value hovers

[3] In conversation with the President of the Parking Lot Attendants Union (Transport Workers Union, Local 700), it was indicated that the cost function was probably linear. This was corroborated in interviews with several enterprisers.

TABLE I—Concentration of Ownership in Terms of
Percentage of Total Capacity Owned by One,
Two and Three Firms: 1955[*]

Market Area	One Firm	Two Firms	Three Firms
1	32	55	66
2	53	76	..
3	5	..	..
4	41	63	86
5	100	..	..
6	52	..	..
7	25	41	..
8	36	..	..
9	45	63	73
10	16	..	..
11	37	51	63
12	26	35	..

[*] Only one or two firms are listed in cases where the addition of the percentage owned by the next largest firm or firms is so small that it was considered insignificant.

Data for the years 1950 to 1954 inclusive may be found in the unpublished dissertaion on file at the University of Pennsylvania, p. 135.

around 107 percent. Of the 70 lots observed, ten did not have 100 percent or more capacity usage at least once during the nineteen observations.

There are, in fact, two capacity concepts: the "comfortable" capacity figure computed by the Traffic Engineer's Office and the one

TABLE II—Capacity Utilization: March 1953

Market Area	Number of Lots	Mean Percentage Utilization	Maximum Percentage Utilization
1	3	97	114
2	5	83	116
3	11	85	105
4	6	86	104
5	0	..	..
6	7	82	104
7	11	87	105
8	5	95	105
9	14	87	116
10	1	86	97
11	3	94	114
12	5	84	95

determined by the ability of the attendant to squeeze another car on the lot, with aisle space often being used.[4] The above concept

[4] Since the pattern of parking varies greatly during the day, it might be suggested that highest prices be charged during the period of greatest demand and lowest prices be charged during the period of lowest demand. This would probably result in a more efficient use of space.

of excess capacity is absolute in the sense that it is applied to all lots without indicating different degrees of excess capacity among lots.

Relative Excess Capacity and a Test for Monopoly

A test for monopoly not only involves the concept "excess capacity," but also involves the concept "excess price." In fact, the term "excess price" is defined as one which creates a *relatively* high degree of "excess capacity." "Excess capacity" (in the relative sense) in turn exists for a firm if there is a significantly lower capacity utilization between the firm and other firms.

It should be kept in mind that, though most firms have excess capacity in the absolute sense, they need not have excess capacity in the relative sense. Thus suppose Firm A has an average capacity utilization of 80 percent and Firm B has an average capacity utilization of 92 percent. This would signify, *ceteris paribus*, that Firm A's price is possibly excessive, since it has, relative to B, excess capacity.

But since parking space is a perishable good (e.g., if it is not sold now, the revenue is lost forever), why should Firm A charge an excessive price? The answer lies in the theory of monopoly price. The multi-lot firm might find a situation of excess capacity profitable because, as the number of lots under the firm's control grows in a given market area, the elasticity of demand for the firm's product decreases.

For example, if a firm has no costs and its demand equation is:

$$P = 1 - \frac{q}{100}$$

where p is the price and q is the quantity of cars parked, the firm would maximize its income when q is fifty. This is shown in Table III. From a social point of view, it may be

TABLE III—Price-Quantity-Income Relationship

Price	Quantity	Income
$.70	30	$21
.60	40	24
.50	50	25
.40	60	24
.30	70	21

argued that the public would best be served if the entire capacity (e.g., output when price is zero) of 100 spaces are sold. But since this output does not maximize income the monopolist prefers a price which results in unused or "excess capacity" of fifty spaces.

With this in mind, the degree of excess capacity and the concentration of ownership were tied together. The question which needs to be answered is: Is there a relationship between concentration of ownership and excess capacity? If the question is answered in the affirmative, there is some ground for suspecting the presence of monopoly. Thus, the null-hypothesis tested was: "The high concentration area has an excess capacity which is not significantly different from the low concentration area." Following standard practice the five percent level of significance was used.

The high concentration area had a mean utilization of 84.81 and a variance of 209.69; the low concentration area had a mean utilization of 88.67 and a variance of 58.92. The "t" value[5] of 1.44 exceeds the five percent level so that no monopoly pricing is adduced.

This, of course, does not say that monopoly elements are not present in the parking market. What is indicated is that pricing at the level which could be considered excessive as measured by the degree of underutilization of capacity could not be found. Thus, if the dominant firm or firms were in a position of price leadership, there could be firms which would possess monopoly power and the results would still be non-significant. The fact that excess capacity was general during March, a peak month, indicates that this is not improbable. A possible test for price leadership would be to determine if there is a lag in time between the price changes of the dominant firms in the high concentration area and price changes for the non-dominant firms with the former changes preceding the latter changes. Unfortunately, adequate data to test for this possibility were not available.

The possibility that the monopoly firm may not be willing or able to maximize profits and therefore produce non-significant results is ruled out. Thus considering the institutional matrix, a firm may desire to avoid adverse

public opinion; or it may attempt to minimize entry of other firms; or finally, it may simply feel there is no reason to bother about maximizing profits, preferring Hicks' "quiet life." In interviews with enterprisers of important firms, it was clearly indicated that there was no fear of public opinion. It was indicated, "this is free enterprise." Finally, the enterpriser finds that it does pay to "bother" with maximizing profits since in this particular case the calculation is not overly complex.

The possibility that prices were kept down in the high concentration area to restrict entry is unlikely since the very scarcity of vacant land is an effective bar to entry. Though entry is not rigidly limited by the scarcity of land because of the possibility of constructing garage facilities, the picture is not changed materially because the cost of building a garage above two tiers is prohibitive. Some evidence of this is seen in the period between 1946 to 1955, a period of high parking lot profits according to the tax investigator's office. Garage spaces increased only from 8,625 to 11,484 while parking lot spaces increased from 11,163 to 12,874. Garages were not included in the analysis because data were unavailable. However, garage prices are practically the same as lot prices in the same location. If garage data were included in the analysis the area of high concentration would have a much higher concentration ratio since practically all the garages are owned by two firms.

One other possibility may be present to produce non-significant results. Lot space may be withheld from other uses in order to maximize their long-run value. While this may be possible in a few cases, the fact that the lots included had a life span generally of six years or more indicates that this factor is minor. Six years of foregone profits in another use is a great sacrifice for anyone. Not only is there a great deal of stability in capacity from year to year for the entire center-city area but this stability is characteristic of the different market area from 1950 to 1955. Furthermore, even though leases run from six months to a year, turnover is small.

One other test for monopoly power is conceptually possible but cannot be displayed here. This would be to test the significance of the difference in excess capacity between the largest firm and all other center-city

---

[5] For the standard error formula employed here see Gerhard Tintner, *Mathematics and Statistics for Economists* (New York, New York: Rinehart & Company, Inc., 1957), p. 260.

firms. Because the identity of the individual firms could not be revealed here, no test was made.

Concluding Remarks

The impact of the automobile has changed market structures considerably. Before the common use of the automobile the business firm that had "a good location" often had monopoly power because of an immobile consumer. The change wrought by the automobile has been to increase the consumers' ability to substitute because of increased mobility. However, this increased use of the automobile has increased the demand for parking space. If entry into the parking market is barred because of lack of space, the problem becomes intensified. The firm that owns most of the lots in a given area assumes a monopoly position. Thus an al-

ready difficult problem is worsened by the possibility of monopolistic pricing. High prices for parking may thus be a factor explaining the exodus to the suburbs and the development of shopping centers.

The significance of the above model lies in its challenge to those who claim that monopoly power cannot be measured. It has been shown that "excess capacity" and "excess price" are functionally related and are measurable. These two factors have been theoretically accepted as indicators of monopoly. It is hoped that this paper not only demonstrates an objective approach to monopoly but also a future area for empirical research.

MYRON H. ROSS

Knox College,
Galesburg, Illinois

Note on the Economics of Residential Zoning and Urban Renewal†

THE proposition that the completely unrestricted use of land (especially urban land) by its owners according to what each deemed to be in his interest would lead to results injurious to all of them as a group has long been accepted as an essential foundation of public policy on land use. By and large, acceptance of this proposition is based on the obvious fact that some forms of land use in any circumstances, and most forms of land use in some circumstances, have beneficial or harmful effects on neighboring properties. As the owner pursuing his own interest will not take these effects into account, the cumulative effect of many owners' unrestricted decisions will not be conducive to their best interests as a group.

An elementary (and classical) example is that of the smoke nuisance. Each owner in a residential district can heat his home or apartment building more cheaply if he burns cheap coal in a smoky furnace than if he burns high grade coal in a furnace that does not emit smoke. However, the value of their dwellings for their own occupancy, or as rental properties, may very well be reduced by more than the saving in heating costs if everyone in the district heats as cheaply as

† Research in connection with the subject matter of this article is being carried out under a grant from the Relm Foundation.

possible. If this is true, they will be benefited by allowing their privacy to be invaded by a regulation, applicable to all of them, restricting the amount of soot and ash their chimneys pour into the air.

The particular kind of nuisance with which this note is concerned is the nuisance of people themselves when they live adjacent to other people whose tastes, habits, and incomes are markedly different from their own. In this case the nuisance may be unilateral rather than mutual. It is generally true that people consider it unpleasant to live near groups of people with lower incomes and with tastes and habits "inferior" to their own, while the reverse is sometimes and perhaps generally not the case.

If the nuisance were mutual there would be no grounds for intervention, as a matter of fact, since the members of each group would gravitate into homogenous residential areas with minimum perimeters bordering the other, mutually repelling, groups' areas. A problem of uneconomic land use, given people's tastes in these matters, will arise only when the nuisance is wholly or partially unilateral. Consider the opposite extreme, where members of group X prefer living near group Y to living entirely surrounded by other members of group X, while members of group Y prefer to live entirely surroun-

ded by other members of group Y. Suppose that streets A, B, C, and D are occupied entirely by members of group X, while streets E, F, G, etc. are occupied entirely by members of group Y; and suppose that only the occupants of streets D and E consider themselves affected by their proximity to members of the opposite group. Under the assumed conditions, if people do not anticipate any change, the properties along street D will sell (and rent) at prices higher than those along streets A, B, and C, and the properties along street E will sell (and rent) at prices lower than those along streets F, G, etc. Suppose also, for simplicity, that lot sizes and structure types are the same in both groups of streets. (This assumption is not realistic but makes no essential difference to the substance of the argument.)

Now if in the initial situation a property on street A, B, or C sells for the same price as a comparable property on street F or G, then the price of a property on street D will sell at a price higher than a comparable property on street E, and it would pay the owners along street E to convert (or sell for conversion) to group X occupancy. In fact, such conversions would continue to be profitable until properties entirely within the group X district sold at prices below the prices of properties within the group Y district by an amount equal to the sum of the premium and discount on each type of boundary property. In taking this profit, however, the owners along street E would be inflicting a loss on the owners on both sides, since they would cause the owners on street D to lose their (desired) proximity to group Y and the owners on street F to acquire (undesired) proximity to group X.

In the initial situation in which properties interior to each group sold at the same prices as in the other, if streets D, E, and F were all under the same owner no conversion would in fact be made because the losses on streets E and F would entirely offset the gain on street D. Conversely, if a final equilibrium under divided ownership were reached, as indicated, with markedly different price levels in the interiors of the two areas, then consolidation of ownership along the boundary would make it profitable for the new owner to convert back to the point at which the price levels of the two interiors were equal. Since such a consolidated owner would make a profit in this operation he could offer higher than

market prices for the properties at and between the two locations for the boundary and it would pay the existing owners to sell to a consolidated owner.

However, there would be a bargaining problem as to how the profit was to be divided between the existing owners and the "re-developer;" and inconsistent expectations could prevent even a mutually profitable bargain from being made. (Such failures are said to be common in real estate experience.) Consequently, there is a case for some form of governmental intervention to maintain or obtain the optimal result: either zoning to keep group X from spreading beyond street D, or a "redevelopment" scheme backed by powers of compulsory acquisition if the over-spill has already occurred.

Holding the line at street D is referred to as "optimal" because it implies equal property values (for properties strictly comparable except for type of occupant) within the two areas. If properties within the area occupied by group Y sell above the prices of comparable properties within the area occupied by group X, there is economic waste in the use of the properties; conversion of some of the properties from the less valuable to the more valuable use will secure the increased value at no cost (or at the cost of the conversion). Generally speaking, it is wise policy to employ resources in their most valuable uses since this is equivalent to maximizing the total product obtained from the resources. Following such a policy consistently has an effect for society as a whole comparable to that obtained by landlords in an area who submit to a uniform restriction on the amount of soot and ash their chimneys emit: it tends to raise the general standard of living through maximizing the total net product even though it may be disadvantageous to some individuals in each particular instance.

As has already been indicated, it will not be profitable to convert any properties (move the boundary between the two groups) for a single owner of all the properties if the prices *at a distance* from the boundary are the same on both sides of the boundary. Suppose the price for a standard property within group X (along streets A, B, and C) is P_x, that the price on the X side of the boundary (along street D) is $P_x + p$; that the price within group Y is P_y, and that the price on the Y side of the boundary is $P_y - d$. If the single owner of streets D, E, and F converts street E to

occupancy by group X, he will gain on that street the difference between the two boundary prices; i.e., his profit on street E will be $(P_x+p)-(P_y-d)$, since he has converted the street from being on the Y side of the boundary to being on the X side. However, on street D he will lose p on each property, because street D is no longer on the boundary but has become interior to group X; and on street F he will lose d on each property, because street F is no longer interior to group Y but has become the boundary of the group. Hence his total profit on a cross-section of single properties, one on each street, will be $(P_x+p)-(P_y-d)-p-d = P_x-P_y$. If the two interior prices, P_x and P_y are equal, this total profit will be zero.

More generally, if the types of dwellings and lot sizes characteristic of groups X and Y are not the same, there will be some cost to converting properties from use by one group to use by the other. In this case a single owner of all the boundary and near-boundary properties would not find it profitable to convert unless the price difference between two typical groups of properties, one from each side (each occupying the same total area) is sufficient to cover the cost of conversion. Under divided ownership, on the other hand, occupancy by group X will in this case expand till the boundary price of group X property exceeds the boundary price of group Y property by the cost of conversion: $(P_x+p)-(P_y-d)=c$, where c is the cost of conversion (assumed the same in either direction). That is, it will proceed to the point where $P_x-P_y=c-p-d$. It will pay a single owner to convert part way back, which is to say that "renewal" will be good public policy, if p+d is greater than 2c; for if so, his profit from renewal, $P_y-P_x-c=p+d-2c$, will be positive. Otherwise, although a wasteful mistake has been made, the gain from undoing any of it is not sufficient to justify the cost.

The economics of urban renewal, or of zoning by residential type, therefore has the following clear (and unexpected) features: (1) There are two ways, equivalent in principle, to measure whether a proposed urban renewal project (or zoning restriction) is economically justified: it is justified if (a) the price of comparable property is higher in the area of preferred use (group Y use in the example) or if (b) the sum of changes in property values at and around the boundary between uses will be positive if conversion is

made toward the preferred use—due allowance being made for costs of conversion in both cases if renewal is contemplated. (2) An "inferior" type of land use, specifically one that causes a nuisance to surrounding properties, should be allowed to expand whenever the price level of properties within the area of that use is higher than the price level of comparable properties within areas of other uses *provided* that the intensity of the nuisance depends only on the distance from the boundary. (This assumption seems appropriate for the type of case under consideration.)

This latter result is especially surprising since we have become quite accustomed to the adoption of renewal projects and of zoning restrictions in the face of substantially higher property values in the inferior use (i.e., slums) than outside it. Clearly the course of policy has not in practice been guided by the considerations mentioned here but has involved other considerations as well.

It is evident from writings and public discussion on this subject that certain types of residential land use, i.e. slums, are felt by most people to be intrinsically undesirable, and that it is thought possible to get rid of them by legislation (zoning) or by demolition and "renewal." Several distinct issues are usually mixed indiscriminately in these discussions and in the present context it will be necessary to separate them. In the first place, people object to the comparative poverty, the low incomes, of most people living in slums. The housing in which they live is largely a reflection of these low incomes although, since there is variation among families as to the amounts they spend on housing and as to the distances they travel to work, it is in part a reflection of differences in tastes in these matters. In the second place, it has been noted that slums and their occupants tend on the average to impose disproportionate tangible costs on the rest of the community because of relatively high morbidity rates (part of which is paid for through public assistance), high crime and delinquency rates, and high costs of fire protection relative to ax revenues.

It is sometimes argued and frequently implied indirectly, that slum clearance would by itself solve these problems although, except in the case of fire risk, there is no relevant evidence indicating that it would. Simple comparisons between morbidity and delinquency rates in public housing and in

the slums from which the occupants came make no allowance for the effects of administrative selection among applicants for public housing. A carefully controlled comparison between occupants and non-occupants of public housing is now in progress in Baltimore, the preliminary results of which indicate that improved housing has no effect on morbidity, i.e. that the observed differences are due entirely to selection.[1] The results on other characteristics are still forthcoming.

Whatever the case for public housing, however, urban renewal is quite another matter, as is restrictive zoning. The effect of these is either to restrict the total available supply of slum housing, holding up prices and rents in the slum areas compared to other areas, or else merely to relocate the slums. Some occupants of slums, at the economically higher margin, may thereby be induced to live in somewhat higher quality housing than would otherwise be the case; and it might be argued that the public will receive benefits from this similar to those, if any, obtained from public (subsidized) housing though to a lesser degree. Even if further study of public housing reveals substantial external benefits, however, this would not imply that such benefits can be of much importance in the analysis of urban renewal or restrictive zoning. They would arise in connection only with a few marginal families who because they are marginal probably contribute relatively little to the external diseconomies of slums in the first place. Certainly, urban renewal and zoning do nothing to relieve the poverty of people in slums, marginal or otherwise.

Indeed, in the light of the remarks made above concerning the economic criteria for optimal zoning and renewal decisions, it can be said that when these activities hold property values and rents higher in slums than elsewhere they lower the real income of the community and therefore, if anything, increase rather than reduce poverty. It would seem much more sensible, therefore, to adopt programs which alleviate poverty, the fundamental problem, rather than attacking the symptom of poor housing.

The one exception to the above line of reasoning has to do with the costs of fire

[1] Daniel M. Wilner, Rosabelle Price Walkley, Marvin N. Glasser, and Matthew Tayback, "The Effect of Housing Quality on Morbidity," Preliminary Findings of the Johns Hopkins Longitudinal Study, *American Journal of Public Health*, December 1958.

protection. There is no reasonable doubt that statistics showing higher costs of fire protection per dollar of assessed value of property can be taken at face value. This implies that the protection of slum properties is being subsidized by other property owners. It does not follow, however, that either zoning or urban renewal is the appropriate means of attack on this problem especially since renewal is in practice much more costly than the subsidy to fire protection in slums. The subsidy would disappear if legislation were enacted imposing a surcharge (extra property tax) on inflammable properties corresponding to the extra cost of their protection. This surcharge would in the short run come out of the net incomes of the owners since their ability to raise rents is already being exploited as far as the traffic will bear. If paying the surcharge was more costly than making the buildings fire-resistant the owners would find it profitable to do the latter and would thereby make the economically efficient decision.

A further possible reason for the existing public support for urban renewal projects and zoning, at points where it is uneconomic, is that people do not like to move even if they are able to convert their properties to slum occupancy (or sell their properties to someone who will convert them) at a profit. Indeed, they may not realize in many cases that there is such a profit; and those who do know but prefer to stay may succeed in getting a project adopted because of the ignorance of the others on this point. If those who do not want to move were really numerically dominant in a neighborhood, however, this would be reflected in property values, and the observed disparities would not exist. Insofar as this motive supplies political force for the adoption of renewal and zoning, therefore these forms of interference in land use represent the subordination of policy to a special interest group at the expense of the community as a whole.

Certain instances exist, however, where the economic criteria developed in this article are in fact satisfied even when slums generally have been prevented from expanding to the optimal extent. Although alternative residential users of property may not be unduly harmed by the necessity of moving, certain institutions specific to a relatively high income neighborhood may find that the costs of moving are prohibitive. In this case prop-

erty values, including the value of the institutional property itself, will in a sense reflect the cost of moving and will accordingly dictate the prevention of the conversion of the neighborhood into slums. If this is true, as might be the case with a hospital or university dependent on a surrounding high-income neighborhood, it would in fact pay the institution to buy up the property surrounding it *even at slum property prices* and to maintain an "island" of the high income, less profitable use. The cost of doing this would, by hypothesis, be less than the cost of moving the institution. Public assistance might be justified on two counts: powers of eminent domain are necessary for the assembly of property in this way because of bargaining problems if the powers are not available; and it may be felt that the institution is deserving of a subsidy.

However, insofar as urban renewal in some cases is in fact a form of federal aid to hospitals and universities, it should be recognized as such and allocated according to principles governing such aid generally. Powers of eminent domain can of course be granted without being accompanied by a subsidy.

The sober appraisal of actual and proposed urban renewal projects requires better factual information than is readily available on the effects of slums and other "inferior" uses on surrounding property values. Real estate experts are not highly communicative on this subject, mainly no doubt because of the variety and complexity of these and other influences on property values. The present writer is currently engaged in an attempt to sort out these factors as reflected in real estate transaction prices in selected areas of Chicago's South Side in a recent ten-year period. If successful, this study may be broadened to other areas. This should in principle make it possible to estimate the external benefits of urban renewal by either of the two formulas, the one relating to boundary values and the one relating to "interior" values.

MARTIN J. BAILEY

Department of Economics,
The University of Chicago

What Ownership Characteristics are Useful in Predicting Response to Forestry Programs?

OWNERS of small forest properties are the target of many types of public and private forestry programs designed to encourage them to adopt more intensive forestry practices. These programs could be designed more effectively if the designers could predict how the owners in a particular area would respond to various types of programs. Although fraught with difficulties, this problem is not unsolvable. For just as foresters can classify stands (by measuring height, age, density, and so on) according to expected response to a thinning treatment, one can probably group forest landowners into "responsiveness" categories based upon easily recognizable owner characteristics.

To do this, the first step would be to find what characteristics are (1) both easy to recognize and measure and (2), closely correlated with owner responsiveness to various types of programs. Several characteristics have been thought to meet both of these requirements. These include occupation of owner, age of owner, method by which forest property was acquired, years owned, distance of forest property from owner's residence, acreage of ownership, and value per acre of the standing timber.

The Northeastern Forest Experiment Station recently conducted a study in southern New York to explore the relationship of such factors to the actual response of owners to various forestry assistance programs.

Sample Based on Adoption or Non-Adoption of Specific Practices. The study sample was drawn to include an adequate sample of both forestry-practice adopters and non-adopters. (Adopters are those individuals who had actually applied one of the specific forestry practices to part of their forest land during the preceding 5 years). And because we thought that the characteristics of owners who planted and those who practiced stand-improvement measures might differ, these categories also were considered in selecting the owners to be visited.

We wanted to know also whether any relationships discovered in one locality would hold true in another locality where market conditions and other factors were somewhat

TABLE I—Relationship Between Ownership Characteristics and Adoption of Forestry Practices for Two Areas in Southern New York: 1958

Practices adopted	Basis: No. of owners	Occupation of Owner				Forest area owned	Timber value per acre	Years owned	Owner's age	Acquired by purchase	Distance to residence
		Farmer	Wood user	Re-tired	Other						
		%	%	%	%	Acres	Dollars	Years	Years	%	Miles
CATTARAUGUS AND CHAUTAUQUA COUNTIES											
Stand-improvement.....	21	38	5	14	43	109	76	13	48	95	8
Planting.........	22	27	0	18	55	91	45	14	50	86	8
None...........	13	38	8	23	31	59	88	18	53	69	0
TIOGA COUNTY											
Stand-improvement.....	11	18	9	0	73	163	38	17	46	81	4
Planting.........	27	30	11	15	44	110	58	18	51	73	6
None...........	20	30	0	25	45	50	28	16	49	79	2

This table is read by making vertical comparisons. For example, do "adopters," on the average, own more forest land than non-adopters? The table shows that in Cattaraugus and Chautauqua Counties forest holdings average 109 acres for those who have done timber-stand-improvement work, 91 acres for those who have planted, and 59 acres for those who have adopted neither practice. Analogous figures for Tioga County are 163 acres, 110 acres, and 50 acres. Thus the answer is Yes, tracts owned by adopters do average considerably larger than tracts owned by non-adopters.

In the case of occupation, the table shows that in Cattaraugus and Chautauqua Counties, 38 percent of those who have done stand-improvement work are farmers, 27 percent of those who have planted are farmers, etc. (not that 38 percent of the farmers have done stand-improvement work.)

different. So we drew our samples from two areas, first from Tioga County, and then from Cattaraugus and Chautauqua Counties. Markets for forest products are somewhat better in the latter area but service foresters and the benefits of other forestry programs of the New York Conservation Department had been available to residents of both areas for a number of years.

Forest Acreage Only Characteristic Related to Forestry Practices Adopted. The basic results of the study are presented in Table I. Most of the ownership characteristics of adopters seem to be essentially the same as those of the non-adopters. In fact, the only statistically significant difference occurred for the classification "acreage of forest land owned." Forest acreage owned by individual adopters averaged considerably greater than forest acreage owned by individual non-adopters. We were rather surprised at the results. We had expected that more of the characteristics

would be useful in identifying owners who respond to forestry programs. However, another phase of the study gave similar results with the owner-occupation characteristic.

Use of Service Not Related to Owner-Occupation. In Tioga County in 1956 and 1957 an owner could do his own planting without any cash outlay and in fact could earn cash income if ACP payments were received. Or, if he preferred, he could give up cash for greater convenience by having his planting done by special crews arranged for by the Tioga County Soil Conservation District. The cost of planting by these crews was $16.50 per acre.

We thought that financial assistance was more likely to bring a response from farmers, who usually have low cash incomes; and that the convenience of having the whole job carried out would appeal particularly to the non-farmers. But, in Tioga County, farmers evidently are as busy as non-farmers—and also are as willing and able to pay for the

planting service. For, as can be seen in Table II, the proportion of tree-planting landowners who used the service was essentially the same for farmers as for non-farmers.

TABLE II—FARMER vs. NON-FARMER USE OF PLANTING SERVICE, TIOGA COUNTY, NEW YORK: 1956-57

Owner class	Owners receiving planting stock	Owners using planting service	Planting owners who use service
1956	Number	Number	Percent
Active farmers........	18	6	33
Others..............	47	12	25
Not classified.........	4	..	..
1957			
Active farmers........	27	5	19
Others..............	66	20	30
Not classified.........	5	..	..

Source: Records and information from Soil Conservation Service, United States Department of Agriculture, Owego, New York, 1958.

Adopters Have Greater Assets. To go back to the one significant characteristic—area of forest land owned—why does this differ between adopters and non-adopters? Is area important in itself or is something more fundamental reflected? We thought that size of tract was simply a measure of the owner's assets and that assets and adoption of forestry practices were positively correlated. We then tested this hypothesis.

Believing a direct measurement of assets was not feasible, we measured them indirectly using the total equalized value of property held by the particular owner in the sample county. On this basis, assets of adopters averaged considerably more than assets of non-adopters (Table III). The differences for one locality (Tioga County) were significant statistically at the 0.05 level and for the other (Cattaraugus and Chautauqua Counties) at the 0.15 level.

The smaller difference between adopters and non-adopters in Cattaraugus and Chautauqua Counties is perhaps more indicative of the difficulties of measuring assets by real estate assessments than it is of any real difference between the two localities. In these two counties nearly twenty percent of the adopters sampled were non-residents,

while in Tioga County the comparable figure was only seven percent. Hence a considerably greater underestimate of the assets of adopters might be expected in

TABLE III—RELATIONSHIP BETWEEN OWNER'S ASSETS AND ADOPTION OF FORESTRY PRACTICES: 1958

Practices adopted	Average assets per owner
Cattaraugus and Chautauqua Counties	
Stand improvement.........	$7,306
Planting...................	6,056
None......................	3,475
Tioga County	
Stand improvement.........	$12,395
Planting...................	9,397
None......................	4,470

Source: County tax assessors' records, total 1957 equalized assessed value of owner's property in the county.

Cattaraugus and Chautauqua Counties. This idea is confirmed when the comparison is limited to residents of the two counties. For these forest landowners, differences in assets between adopters and non-adopters are very nearly significant at the 0.05 level.

Conclusions

Those responsible for developing and administering forestry programs would like to design their programs to obtain as great a response from forest landowners as possible. Knowing something about these landowners should help in developing programs to which they will respond. But what is it that we want to know about these people?

The results of this study cast considerable doubt on the usefulness of several commonly used types of information about forest landowners. For indicating the probable response to forestry programs in southern New York only two ownership characteristics tested in this study would appear to be particularly helpful. They are "acreage of forest-land owned" and "assessed value of the owner's property." Both of these variables are positively correlated with response to public programs offering assistance to landowners in planting and in doing stand improvement work.

This implies that, if timber production at minimum cost is the objective, forestry as-

sistance efforts might best be concentrated on the more well-to-do landowners. To do this would conflict with the use of forestry programs as income-assistance measures. Thus the necessity of clearly specifying the objectives of a particular program is underscored.

These results offer hope that easily recognizable ownership characteristics can be found that will be of help in planning more effective forestry programs. But the results also point to the importance of testing the usefulness of various kinds of information before spending considerable sums to collect it.

Additional ownership studies are needed to test the relationship between easily recognizable ownership characteristics and response to various types of forestry programs.

Surveys designed primarily to collect information on the characteristics of owners in an area would logically follow the ownership studies; they would be conducted after the studies indicate what characteristics are helpful.

HENRY H. WEBSTER
Forest Economist

CARL H. STOLTENBERG,
Chief, Division of Forest
Economics Research

Northeastern Forest Experiment Station
Forest Service, United States
Department of Agriculture
Upper Darby, Pa.